Working Paper 98-21
Statistics and Econometrics Series 14
February 1998

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

LOCAL CROSS VALIDATION FOR SPECTRUM BANDWIDTH CHOICE.

Carlos Velasco*

Abstract

We investigate an automatic method of determining a local bandwidth for nonparametric kernel spectral density estimates at a single frequency. This procedure is a modification of a cross-validation tecnique for global bandwidth choices, avoiding the computation of any pilot estimate based on initial bandwidths or on approximate parametric models. Only local conditions on the spectral density around the frequency of interest are assumed. We illustrate with a Monte Carlo study the performance in finite samples of the bandwidth estimates proposed.

Keywords:
Bandwidth selection, nonparametric spectral estimation, cross-validation, time series, periodogram.

*Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 126 28903 Madrid. Spain. Ph: 34-1-624.98.87, Fax: 34-1-624.98.49, e-mail: cavelas@est-econ.uc3m.es. This paper is a revised version of the Chapter 4 of the author´s University of London Ph.D. Thesis. I am grateful to Professor P.M. Robinson for his guidance and advice. Research funded by the Spanish Dirección General de Enseñanza Superior, Ref. n. PB95-0292.

Local Cross Validation for Spectrum Bandwidth Choice

Carlos Velasco

Departamento de Estadística y Econometría

Universidad Carlos III de Madrid

Calle Madrid 126

28903 Getafe (Madrid)
Spain

Local Cross Validation for Spectrum Bandwidth Choice

Carlos Velasco*
Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid 126
28903 Getafe (Madrid)
Spain

February 2, 1998

Abstract

We investigate an automatic method of determining a local bandwidth for nonparametric kernel spectral density estimates at a single frequency. This procecdure is a modification of a cross-validation technique for global bandwidth choices, avoiding the computation of any pilot estimate based on initial bandwidths or on approximate parametric models. Only local conditions on the spectral density around the frequency of interest are assumed. We illustrate with a Monte Carlo study the performance in finite samples of the bandwidth estimates proposed.

Keywords. Bandwidth selection; nonparametric spectral estimation; cross-validation; time series; periodogram.

1 Introduction

Smoothed estimation of the spectral density of stationary time series, like many nonparametric methods of inference, relies on the choice of a bandwidth or lag number depending on the sample size. The properties of the estimates depend crucially on the value of this number. Asymptotic theory prescribes a rate for the lag number M with respect to the sample size N as this tends to infinity, but gives no practical guidance for the choice of M in finite samples. Different techniques have been proposed in the literature to that end. The usual criterion is the minimization of some

*This paper is a revised version of the Chapter 4 of the author's University of London PhD Thesis. I am grateful to Professor P.M. Robinson for his guidance and advice.

estimate of the asymptotic mean square error of the estimator. This can be implemented by plug-in or cross-validation methods. Also, global and local choices are possible, depending on whether we are interested in the behaviour of the spectral density for all range of frequencies or in a specific point or small interval.

The plug-in method consists in substituting the unknowns of the leading term in the asymptotic expression for the mean square error by consistent estimates, generally nonparametric, but also parametric ones based on approximate models can be used. Cross validation procedures avoid the use of those initial estimates and approximate the mean square error indirectly. They are based on estimates which do not use the information contained in the sample about the function of interest at each point (at each Fourier frequency in the case of spectral estimation). Wahba (1980) considered automatic smoothing methods for the log periodogram, but in many cases we are interested in obtaining optimal bandwidths for the original scale.

Beltrao and Bloomfield (1987) (BB hereafter) considered bandwidth choice for discrete periodogram average type spectral estimates. They justified a method based on a cross-validated form of Whittle's frequency domain approximation to the likelihood function of a stationary Gaussian process (see also Hurvich (1985)). Robinson (1991) extended their results under more general conditions for a wider class of models, including spectral estimation for the construction of efficient regression estimates, and proved the consistency of the estimate of M. This cross-validated method selects a global bandwidth for all the range of frequencies $[-\pi, \pi]$ or for a fixed subset of it. Here we propose a modified version of cross validation to justify a local bandwidth choice for a single frequency, following some ideas suggested in Robinson (1991, p. 1346), related with the work of Hurvich and Beltrao (1994) in a different context. For this single frequency choice, we only use local smoothness properties of the spectral density of the time series around this frequency, allowing for a broader range of dependence models. This local adaptation could lead also to efficiency gains when estimation of the spectrum for all range of frequencies $[-\pi, \pi]$ is in mind.

The method we analyze here can be seen as the cross validation alternative to Bühlmann's (1996) iterative local plug-in procedure for lag-window spectral estimates, proposed by Brockmann et al. (1993) in the context of kernel regression estimators (see also Herrmann (1997)), or to the related proposal of Newey and West (1994) for covariance matrix estimation. Local adaptation is also studied by Lepskii and Spokoiny (1995) for projective estimates in a "signal+noise" model. Here the range of estimation is split on degenerating intervals with the asymptotics and different smoothing parameters are estimated independently for each one.

Next section is devoted to the assumptions that we will use in this paper, together with a brief introduction to the main cross validation concepts for nonparametric spectrum estimation and a detailed analysis of the mean square error for the spectral estimate at a fixed frequency under local smoothness assumptions. Section 3 introduces the local cross validation criterion and the main result of the paper. Then we carry out a Monte Carlo analysis of the finite sample behaviour of the techniques proposed. All the proofs and some technical lemmas required are given in the Appendix.

2 Assumptions and definitions

In this section we will introduce some assumptions and definitions, together with some intuitions about cross validation and BB's results. Given the observed data X_t, $t = 1, 2, \ldots, N$, the periodogram at the frequency $\lambda_j = 2\pi j/N$, j integer, is equal to

$$I(\lambda_j) \stackrel{def}{=} \frac{1}{2\pi N}\left|\sum_{t=1}^{N} X_t \exp\{it\lambda_j\}\right|^2 .$$

The averaged-periodogram spectral estimate with lag number $M = M_N = h_N^{-1}$, where h_N is the bandwidth of the estimate in BB's notation, and kernel or spectral window K (this function was denoted by W in BB, but we use this notation later for another analogous function), is

$$\widehat{f}_M(\lambda_j) \stackrel{def}{=} \sigma_M^{-1} \sum_k K(M\lambda_k)\, I(\lambda_j - \lambda_k),$$

where the summation runs for all values of k in the support of K (not including all the periodogram ordinates equal to the zero frequency periodogram to account for mean correction). and σ_M gives the exact sum of the weights used

$$\sigma_M \stackrel{def}{=} \sum_k K(M\lambda_k).$$

Here we could have used the value $2\pi M/N$ instead of σ_M^{-1}, using that K integrates to 1, but this simplifies some arguments. We stress the dependence of $\widehat{f}_M$ on M in the notation, since this is the *parameter* of interest.

BB (cf. their Theorem 3.1) considered a zero mean stationary Gaussian process $\{X_t\}$ with autocovariance function $\gamma(r) = E[X_0 X_r]$ satisfying

$$\sum_1^{\infty} r|\gamma(r)| < \infty,$$

and spectral density $f(\lambda) = (2\pi)^{-1}\sum_{-\infty}^{\infty}\gamma(r)\exp\{ir\lambda\}$ everywhere positive. The kernel function K they used for the nonparametric estimates was non-negative, even, bounded function, with

$$\int_{-\infty}^{\infty} K(x)dx = 1, \quad \int_{-\infty}^{\infty} x^2 K(x)dx < \infty.$$

Also we can write $K(x) = \int w(y)\exp\{ixy\}dy$, where w is of compact support. The bandwidth h_N satisfies $h_N^{-1} = O(N^\rho)$, for some $\rho < \frac{2}{5}$ and $h_N = o(1)$.

The 'leave-two-out version' of the estimator $\widehat{f}_M$ (we leave only two frequencies out if K were actually compactly supported inside $[-\pi, \pi]$, as we will assume later on, or if we had defined its periodic version in that interval) is

$$\widehat{f}_M^{\mathbf{j}}(\lambda_j) \stackrel{def}{=} \sigma_{j,M}^{-1} \sum_k{}' K(M\lambda_k)\, I(\lambda_j - \lambda_k), \tag{1}$$

where $\sum_k{}'$ runs for the same values as before, except in the set of indices of frequencies $\lambda_j - \lambda_k$ with the same periodogram ordinate as $I(\lambda_j)$, i.e., $k \in \{0, \pm N, \ldots\} \cup \{2j, 2j \pm N, \ldots\}$. Also, the normalizing number $\sigma_{j,M}$ is now equal to

$$\sigma_{j,M} \stackrel{def}{=} \sum_k{}' K(M\lambda_k).$$

Introduce the pseudo log-likelihood type criterion

$$L(f) \stackrel{def}{=} \sum_{j=1}^{N-1} \{\log f(\lambda_j) + I(\lambda_j)/f(\lambda_j)\}, \tag{2}$$

which is Whittle's approximation for the likelihood of a Gaussian sequence in the frequency domain. BB showed under the previous conditions that

$$L\left(\widehat{f}_M\right) - L(f) = \frac{N}{2}\mathrm{IMSE}(M)$$

plus a term of smaller order in probability, where IMSE(M) is the discrete approximation to the Integrated Mean Squared Error of $\widehat{f}_M$, weighted by f^{-1}:

$$\mathrm{IMSE}(M) \stackrel{def}{=} N^{-1} \sum_{j=1}^{N-1} E\left[\left\{\widehat{f}_M(\lambda_j) - f(\lambda_j)\right\}/f(\lambda_j)\right]^2.$$

Then minimizing $L\left(\widehat{f}_M\right)$ and $\mathrm{IMSE}(M)$ should be approximately equivalent, and this is the basis for the estimation of the M that minimizes $\mathrm{IMSE}(M)$ for $\widehat{f}_M(\lambda)$ in $[-\pi, \pi]$.

If we are interested in nonparametric spectral estimation at a single frequency (of special interest is the zero one; see Bühlmann's (1996) examples, together with covariance matrix estimation in econometrics, like in den Haan and Levin (1996) and the references therein) or we want to achieve possible efficiency gains using different bandwidths for each frequency, we need a criterion to choose a local bandwidth. The Mean Square Error at a frequency ν,

$$\mathrm{MSE}(\nu, M) \stackrel{def}{=} E\left[\left\{\widehat{f}_M(\nu) - f(\nu)\right\}/f(\nu)\right]^2,$$

is the usual criterion employed to assess nonparametric estimates of this class at a single frequency. We consider only fixed frequencies of the form $\nu = 2\pi v/N$, where v is an integer. We restrict to $0 \leq v \leq \frac{1}{2}N$, given the symmetry and periodicity of the periodogram and the spectral density. We will use the following Assumptions:

Assumption 1 *X_t, $t = 1, 2, \ldots$ is a Gaussian stationary time series.*

Assumption 2 *The spectral density $f(\lambda)$ of X_t has three uniformly bounded derivatives in an interval around the fixed frequency ν, with $f(\lambda) > 0$ for λ in that interval, and $f \in L_p[-\pi, \pi]$ for some $p > \frac{5}{3}$.*

Assumption 3 *The function K is non-negative, even, bounded, zero outside $[-\pi, \pi]$, of bounded variation and*

$$\int_{-\infty}^{\infty} K(x)dx = 1, \quad \int_{-\infty}^{\infty} x^2 K(x)dx = \omega_2 < \infty.$$

Assumption 4 *The function K has Fourier transform $w(x) = (2\pi)^{-1} \int_{-\infty}^{\infty} K(\lambda) e^{i\lambda x} d\lambda$ satisfying*

$$|w(x)| = O(|x|^{-\alpha}) \quad as \quad |x| \to \infty,$$

for some $\alpha > \frac{5}{4}$.

Assumption 1 was used also in BB, but we do not need to assume zero mean since we avoid the zero frequency periodogram ordinate in the definition of our estimates. Assumption 2 only requires smoothness properties of f around the frequency we are interested in, allowing for a wide class of spectral densities, including ones with zeros and poles outside a neighbourhood of ν. The only requirement outside this band is an integrability condition to ensure ergodicity (with respect to second moments) of the series (see Lemma 7 below).

A compact support kernel in Assumption 3 is then the complementary of Assumption 2 in order to guarantee that we only use information in an interval around ν. The rest of conditions on K are standard, Assumption 4 being necessary to approximate $\widehat{f}_M$ with a weighted autocovariance type estimate in Lemma 5. From this lemma, both estimates have the same asymptotic distribution and mean square error, so the bandwidth choice techniques for one are valid for the other. This condition is satisfied by the Barlett-Priestley and quadratic spectral kernels (with $\alpha = 2$), but not by the Daniell or uniform spectral window.

With Assumption 3, the summation in k in the definition of $\widehat{f}_M$ takes values in $\{j - N + 1, \ldots, j-1\} - \{j\}$ due to the compact support kernel, and in $\{j-N+1, \ldots, j-1\} - \{j, 2j, 2j-N\}$ for $\widehat{f}_M^{\,j}(\lambda_j)$.

We now present a result concerning the mean square error of the estimate $\widehat{f}_M$ at Fourier frequencies, which will be used to analyze a local version of the likelihood (2). We use in the proof two lemmas given in the Appendix about the discrete Fourier transform and periodogram of the observed sequence, extending and correcting some of the results of BB, assuming only local smoothness for the spectral density. We have to distinguish between estimates for Fourier frequencies λ_j close to the origin, and at remote frequencies. Define $\|K\|_2^2 = \int K^2(x)dx$ and c a finite positive constant, not necessarily always the same.

Lemma 1 *Under Assumptions 1, 2 and 3, if $M = c \cdot N^{1/5}$, for frequencies $\lambda_j = 2\pi j/N$ such that $|\nu - \lambda_j| \leq c \cdot m^{-1}$ for some positive sequence m such that $1/m + m/M \to 0$, then, uniformly in j, for $\underline{\nu > 0}$,*

$$MSE(\lambda_j, M) = \frac{M}{N} 2\pi \|K\|_2^2 f(\lambda_j)^2 + M^{-4}\left[\frac{w_2}{2} f^{(2)}(\lambda_j)\right]^2 + O\left(\frac{M^2}{N^2} + N^{-1}\right), \tag{3}$$

and for $\underline{\nu = 0}$,

$$MSE(\lambda_j, M) = \frac{M}{N} 2\pi \|K\|_2^2 f(\lambda_j)^2 [1 + \delta_M(j)] + M^{-4}\left[\frac{w_2}{2} f^{(2)}(\lambda_j)\right]^2 + O\left(\frac{M^2}{N^2} + N^{-1}\right), \tag{4}$$

where $0 \leq \delta_M(j) \leq 1$ measures the degree of overlapping between different kernels K at a distant $2M\lambda_j$ apart when $\lambda_j \to 0$ as $N \to \infty$. For $j = 0$, $\delta_M(j) = 1\ \forall M$, and for $\lambda_j > 2\pi/M$, $\delta_M(j) = 0$.

For $\nu > 0$ this is the standard result for globally smooth spectral densities (see for example Brillinger (1975), Corollaries 5.6.1 and 5.6.2). However in a degenerating band around the origin (small λ_j), the nonparametric spectral estimates have variance depending on the overlapping of two kernel functions K centred at frequencies λ_j and $-\lambda_j$ respectively, measured by the quantity $\delta_M(j)$.

To make the bias and the variance of the same order of magnitude we would take $M = \tau N^{1/5}$, for some $0 < \tau < \infty$ and then MSE will be of order $M^{-4} \sim cM/N$. From the previous lemma, the optimal constant $\tau^\star$ that minimizes the leading term of MSE of $\widehat{f}_M(\nu)$ is

$$\tau^\star = \left\{\frac{\omega_2 f^{(2)}(\nu)^2}{2\pi \|K\|_2^2 f(\nu)^2}\right\}^{1/5}, \tag{5}$$

if $\nu \neq 0$ and with 4π instead of 2π for $\nu = 0$. Now it is possible to estimate the value of $\tau^\star$ using initial, pilot estimates of the spectral density and its second derivative at ν. This is the approach of several authors, including Andrews (1991), Newey and West (1994) or Bühlmann (1996), just to give some recent contributions. In the following section we adopt instead an indirect approach using a cross-validation argument.

3 Local cross validation

Consider for some positive sequence $m = m_N$ such that $m^{-1} + m/M \to 0$ as $N \to 0$, one form of *local* integrated mean square mean,

$$\text{IMSE}_m(\nu, M) \stackrel{def}{=} \frac{2\pi}{N} \sum_{j=1}^{N-1} W_m(\lambda_j - \nu) \text{E}\left[\left\{\widehat{f}_M(\lambda_j) - f(\lambda_j)\right\} / f(\lambda_j)\right]^2,$$

where $W_m(\lambda) = m\sum_j W(m[\lambda + 2\pi j])$ for some appropriate kernel function W satisfying Assumption 3. For the uniform kernel $W = (2\pi)^{-1}I_{[-\pi,\pi]}$ and $m = 1$, we have $\mathrm{IMSE}_m(\nu, M) = \mathrm{IMSE}(M)$ for all ν.

Then, from Lemma 1 and $\underline{\nu > 0}$, we can obtain under the same regularity conditions, as m increases with N,

$$\begin{aligned}\mathrm{IMSE}_m(\nu, M) &= \frac{M}{N}2\pi\|K\|_2^2 + M^{-4}\left[\frac{w_2}{2}\frac{f^{(2)}(\nu)}{f(\nu)}\right]^2 + O\left(\frac{M^2}{N^2} + \frac{1}{N} + \frac{M}{N}\left[\frac{m}{N} + \frac{1}{m}\right]\right)\\ &= \mathrm{MSE}(\nu, M) + o(\mathrm{MSE}(\nu, M)),\end{aligned}$$

where the errors in m come from the continuous approximation to the sum in IMSE_m and we use that the ratio $f^{(2)}(\nu)/f(\nu)$ has bounded derivative. Therefore $\mathrm{IMSE}_m(\nu, M)$ approximates $\mathrm{MSE}(\nu, M)$ when $\nu > 0$ as $m \to \infty$.

When $\underline{\nu = 0}$, we can see that

$$\begin{aligned}\mathrm{IMSE}_m(0, M) &= \frac{2\pi}{N}\sum_{j=1}^{N-1} W_m(\lambda_j)\left[\frac{M}{N}2\pi\|K\|_2^2\{1 + \delta_M(j)\}\right]\\ &\quad + M^{-4}\left[\frac{w_2}{2}\frac{f^{(2)}(0)}{f(0)}\right]^2 + O\left(\frac{M^2}{N^2} + N^{-1} + \frac{M}{N}\left[\frac{m}{N} + m^{-1}\right]\right).\end{aligned} \tag{6}$$

Now in the summation in (6) we can consider the values of λ_j smaller and bigger than $2\pi/M$ in absolute value. Since $|\delta_M(j)| \le 1\ \forall j$, $|\delta_M(j)| = 0$ if $|\lambda_j| > 2\pi/M$ (i.e. $|j| > N/M$) and $m/M \to 0$, with $\sup_{m,j}|W_m(\lambda_j)| = O(m)$,

$$\begin{aligned}&\frac{2\pi}{N}\sum_j W_m(\lambda_j)\left[\frac{M}{N}2\pi\|K\|_2^2\{1 + \delta_M(j)\}\right]\\ &= \frac{M}{N}2\pi\|K\|_2^2\frac{2\pi}{N}\sum_j W_m(\lambda_j) + \frac{M}{N}2\pi\|K\|_2^2\frac{2\pi}{N}\sum_{|j|\le N/M} W_m(\lambda_j)\delta_M(j) &(7)\\ &= \frac{M}{N}2\pi\|K\|_2^2\frac{2\pi}{N}\sum_j W_m(\lambda_j) + O\left(\frac{m}{N}\right)\\ &= \frac{M}{N}2\pi\|K\|_2^2 + O\left(\frac{M}{N}\frac{m}{N}\right) + o\left(\frac{M}{N}\right). &(8)\end{aligned}$$

Therefore, when $\nu = 0$, the quantity $\mathrm{IMSE}_m(0, M)$ only estimates half of the asymptotic variance in $\mathrm{MSE}(0, M)$, though the second term in (7), corresponding to the overlapping factor in Lemma 1, of magnitude m/N, will contribute to $\mathrm{IMSE}_m(0, M)$ in finite samples.

A possible approach to obtain a consistent estimate of the optimal local bandwidth which minimizes $\mathrm{MSE}(\nu, M)$, $M^* = \tau^* N^{1/5}$, is to minimize an estimate of $\mathrm{MSE}(\nu, M)$ or of $\mathrm{IMSE}_m(\nu, M)$, which approaches the former as m increases. Some adjustments might be necessary in the case $\nu = 0$ due to the problem described in the previous paragraph. The presence of two related bandwidth parameters, m and M, seems to imply a circular argument like the one present in a plug-in method, where pilot estimates of the spectral density and its derivatives are used, depending on other bandwidths or parametric assumptions. To circumvent this problem we describe some procedures in the next section that connect both choices, showing that the choice of m is not too decisive.

The logical cross validation argument in this case would be the minimization with respect to M of the function (recalling the definition of the 'leave-two-out' spectral estimate in (1)),

$$\mathrm{CVLL}_m(\nu, M) \stackrel{def}{=} 2\pi \sum_{j=1}^{N-1} W_m(\lambda_j - \nu)\left\{\log \widehat{f}_M^{\mathbf{j}}(\lambda_j) + I(\lambda_j)/\widehat{f}_M^{\mathbf{j}}(\lambda_j)\right\},$$

which is a likelihood that tends to use only the information around ν as $m \to \infty$. Since W has compact support $[-\pi, \pi]$, just about N/m frequencies around ν are used. It is likely that this procedure leads to more variability than the global one, since we are not using all information of the sample (see Bühlmann (1996), Section 3.1, or Brockmann et al. (1993) for a related problem in nonparametric regression).

To justify the above ideas we have the following Proposition, proved in the Appendix.

Proposition 1 *Under the Assumptions 1, 2, 3, 4, W satisfying Assumption 3, $M = c \cdot N^{1/5}$ and $m^{-1} + m/M \to 0$,*

$$\begin{aligned} CVLL_m(\nu, M) &= 2\pi \sum_{j=1}^{N-1} W_m(\lambda_j - \nu)\left\{\log f(\lambda_j) + I(\lambda_j)/f(\lambda_j)\right\} \\ &\quad + \frac{N}{2}\, IMSE_m(\nu, M) + o_P(N\, IMSE_m), \end{aligned}$$

where $0 < c_1 < IMSE_m/IMSE < c_2 < \infty$ as $N \to \infty$, and the first term on the right hand side depends only on m (but not on M).

Then, under regularity conditions, $\mathrm{CVLL}_m(\nu, M)$ is a consistent estimator of $\mathrm{IMSE}_m(\nu, M)$ up to a constant not depending on M. From there, minimization of CVLL_m should be approximately equal to minimization of IMSE_m. Since the latter approximates $\mathrm{MSE}(\nu)$ under similar conditions on m, we can expect to obtain reasonable estimates of the local optimal M using the local cross-validation criterion with $\widehat{M}(\nu) = \arg\min_M \mathrm{CVLL}_m(\nu, M)$.

BB did not require to estimate explicitly IMSE or its asymptotic rate of convergence, but in our case we need to do so because we estimate a local MSE from an IMSE calculated from estimates around the frequency of interest. To this end, additional stronger conditions are required for the spectrum at that frequency, but we do not need to make global assumptions for the spectral density.

4 Monte Carlo work

In this section we assess if all the asymptotic arguments given in previous sections are good approximations for reasonable finite sample sizes and whether the cross-validation leads to sensible bandwidth estimations. We have concentrated first on the special case of the estimation of the bandwidth for nonparametric spectral estimates at the origin ($\nu = 0$) and then on the estimation of the spectral density for all $\lambda \in [-\pi, \pi]$, following Bühlmann's (1996) Section 3.

We have simulated Gaussian sequences following five different models and sample sizes N=120, 256 and 480. The models considered are the following AR processes,

$$X_t = \sum_{j=1}^{p} \alpha_j X_{t-j} + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0,1),$$

with parameters

MODEL 1, AR(3): $\alpha_1 = 0.6, \quad \alpha_2 = -0.6, \quad \alpha_3 = 0.3.$

MODEL 2, AR(2): $\alpha_1 = 0.6, \quad \alpha_2 = -0.9.$

MODEL 3, AR(1): $\alpha_1 = 0.8.$

MODEL 4, AR(2): $\alpha_1 = 1.372, \quad \alpha_2 = -0.677.$

MODEL 5, AR(5): $\alpha_1 = 0.9, \quad \alpha_2 = -0.4, \quad \alpha_3 = 0.3, \quad \alpha_4 = -0.5, \quad \alpha_5 = 0.3.$

and $\alpha_i = 0$ if not stated. The last three parameter sets were used also by Bühlmann (1996). These models are convenient because of their simplicity and the different spectra they represent. From Figure 1, Model 1's spectral density exhibits a small peak at the origin and a larger one at $\lambda \approx 1.5$. Model 2 is flat at the origin, but with a very sharp peak at frequency $\lambda \approx 1.3$. The AR(1) Model 3 has the typical spectral density of an AR(1) series with positive autocorrelation and a maximum at zero frequency. The AR(2) spectrum of Model 4 is similar to the first one, but with a minimum at the origin and a closer peak, whereas Model 5's spectrum shows several peaks, including one at origin.

With these processes we hope to assess the performance of the approximations in situations where global bandwidths might be not be very appropriate due to the presence of special features in the spectral density at the frequency of interest or at remote frequencies which can distort global procedures.

We have not normalized the time series to have equal variance or same spectral density at the origin, since this would only imply multiplying the periodogram of the observed time series by a fixed constant and will not affect any of the procedures used.

For the local choice at $\lambda = 0$, we employ the Barlett-Priestley Kernel (for both K and W), with spectral window

$$K(\lambda) = \begin{cases} \frac{3}{4\pi}\left\{1 - \left(\frac{\lambda}{\pi}\right)^2\right\}, & |\lambda| \leq \pi, \\ 0 & |\lambda| \geq \pi \end{cases}$$

and lag-window

$$w(x) = \frac{3}{(\pi x)^2}\left(\frac{\sin \pi x}{\pi x} - \cos \pi x\right).$$

The uniform kernel was also tried for K, with much less smooth results as a consequence of the non-continuity in the boundaries of its support and a lag-window with tails slowly decreasing to zero. For the choice at all frequencies $\lambda_j \in [-\pi, \pi]$ we report the results for W equal to the uniform kernel, in this case not being very different from those with the Barlett-Priestley window.

The tables with the simulation outcomes and the plots are given at the end of the paper.

4.1 Spectral estimation at the origin

From equation (8) we know that for the frequency $\nu = 0$ in particular, $\text{IMSE}_m(0, M)$ does not approach $\text{MSE}(0, M)$ asymptotically due to the different variance of the spectral density estimates around the origin. Nevertheless, from Lemma 1, the transition from the variance of $\widehat{f}_M(0)$ to the variance of an estimate at a frequency apart from the origin (one half of the previous one) is smooth, depending on the shape of the kernel used. Then we can expect that the approximation behaves moderately well also for this case.

We have used the following equivalent version of the cross-validated log-likelihood, given the periodicity and symmetry of W_m, f and I,

$$\text{CVLL}_m^\star(0, M) \stackrel{def}{=} 2\pi \sum_{j=-[N/2]}^{[N/2]} W_m(\lambda_j)\left[\log \widehat{f}_M^{\mathbf{j}}(\lambda_j) + I(\lambda_j)/\widehat{f}_M^{\mathbf{j}}(\lambda_j)\right],$$

dropping the frequency $\lambda_j = 0$ (since due to mean correction $I(0) = 0$ and $\widehat{f}_M(0) = \widehat{f}_M^{\mathbf{0}}(0)$) and we define $\text{IMSE}_m^\star$ accordingly.

4.1.1 Results for IMSEm

The first goal is to check if $\text{IMSE}_m(0, M)$ estimates $\text{MSE}(0, M)$ properly and how sensitive is to the choice of m. Specially interesting are the cases with moderate values of m, smaller than the optimal M^*, for which $\text{CVLL}_m(0, M)$ should be estimating $\text{IMSE}_m(0, M)$ according to Proposition 1. Due to the problems commented before we cannot expect high precision at frequency zero, but at least certain information about the shape of the spectral density in intervals around the origin.

To evaluate IMSE_m, we first estimate $\text{MSE}(\lambda_j, M)$ by Monte Carlo (with 1000 replications and sample size $N = 256$) for all j and a grid of $M = 1(0.5)30$, which cover all reasonable M's, including the optimal values for the sample size considered. Then $\text{IMSE}_m(0, M)$ is evaluated for different values of m and the minimum with respect to M found. The values of m were chosen (see Table I) in terms of the number of different Fourier frequencies around λ_0 over which the kernel W averages in each case, denoted as 'band':

$$\text{band} = \frac{N}{2m}.$$

The correspondent grid is band=1(4)129, which covers all the possibilities for $N = 256$. (The optimal values are calculated using the pointwise result (5) for the MSE at a single frequency, $\tau^{\star}$, depending on the kernel used and on the values of $f(0)$ and its second derivative.)

The results are reported in Table I and the correspondent plots are in Figures 2 to 6 (in the two-dimensional graphs each horizontal line corresponds with one value of m). From high values of m we can check that the asymptotic expression for the optimal M for $\widehat{f}_M(0)$ is not very precise for this sample size for most of the models tried. For moderate values of m, up to about 4, the approximation is quite reasonable in most cases, except for Model 2, certainly due to the consideration in IMSE_m of frequencies corresponding to the sharp peak, which leads to large $\widehat{M}$'s to improve the nonparametric estimation there, and for Model 3, where we are trying to estimate a sharp peak itself but the rest of the spectrum is totally flat. Another interesting feature is the stability of the M's minimizing IMSE_m for values of m from 1 to 3, which hopefully will extend to the estimates we propose, based on CVLL_m. This can be checked as well in Figures 2 to 6, where the minimum for each value of m is always in the same range of values of M, except, perhaps, for a small number of lines (each corresponding to a different value of m in Table I).

4.1.2 Results for CVLLm

We next estimate the function $\text{CVLL}_m(0,M)$ for a grid of values of m and M and then we report in Table II the bias, standard deviation and mean square error of the M estimated by the minimization of CVLL_m (1000 replications). The conclusions here are similar to those of IMSE_m. The lag number M estimated on CVLL_m shows a moderate bias and about the same standard deviation for all values of m between 1 and 5. In the case of Model 2, much higher values of $\widehat{M}$ are estimated than the asymptotic optimal ($M^* = 7.13$), agreeing with the IMSE_m findings. Similar observation holds in inverse direction for Model 3, where much smaller values than $M^* = 20.34$ are found. In the bi- and tri-dimensional plots of Figures 7 to 11 we can only give some of the CVLL_m lines due to very different scales.

As expected, when large values of m ($\geq$6) are used in CVLL_m we obtain in some cases much reduced biases for $\widehat{M}$ (including Models 2 and 3), but due to the use of a very small number of spectral estimates, this leads to quite imprecise estimates (high standard deviations). Nevertheless, occasionally these estimates of M based on CVLL_m with high m have smaller mse (which is calculated with respect to M^*, and not with respect the value of M found in the previous subsection looking at the simulated IMSE_m for each m).

Summarizing, we find that CVLL_m reflects the different characteristics of the spectral density for a range of moderate values of m and can be a useful means of studying local properties of the spectral density. The variability of the estimates is relatively high, as in most of bandwidth choice methods (characterized by slow rates of convergence) and like in any nonparametric method this variance tends to increase in general with the value of m (which is proportional to the inverse of the actual bandwidth of the kernel W_m).

4.1.3 Approximation of M

From a theoretical point of view, the choice of m has not a definitive answer, though we have just seen that this might be not too decisive. In practical applications a first possibility is a selection criteria depending only on a fix m, smaller than about 4 for sample size $N = 256$ say, from the previous subsection. For any sample size this would imply to use about N/m Fourier frequencies in CVLL_m. We can also make this choice dependent on N. In Tables III to VII we have tried the following choices for sample sizes $N = 120,\ 256,\ 480$ and Models 1 to 5:

- 'GLOBAL': $m = 1$. This is the same as BB's global procedure.
- '1': $m = 1.4$.
- '2': $m = N^{0.01}$.
- '3': $m = N^{0.02}$.
- '4': $m = N^{0.04}$.

These cover all reasonable values for the three sample sizes in the light of the behaviour of CVLL_m.

To reduce the dependence on this quite arbitrary decision, we can then use the local $\widehat{M}$ estimated in these initial stages to construct a choice of m that adapts also locally to the shape of f. For each of the previous five initial estimates of M we have tried $\widehat{m} = \widehat{M} \cdot N^{-0.03}$, in agreement with Proposition 1. These are contained in the rows labelled '5' to '9'.

Another alternative consists in starting with a fairly global choice of m (i.e. small m), and then iterate the values of M and m successively with the same recursion as before. This type of procedure did not tend to converge always, independently of the initial value of m, so we had to decide how to choose M if the iteration limit, 5, was reached. We used the following alternatives in the case of convergence problems:

- 'ITER.1': $m_0 = 1$ and if not converging we take the $\widehat{M}$ estimated with smallest CVLL_m for the last three m's tried (hoping to achieve a better minimization of the local IMSE).
- 'ITER.2': $m_0 = 1.5$ and we proceed as before.
- 'ITER.3': $m_0 = 1$ and we take the most global choice of the last three if convergence is not achieved, i.e. the $\widehat{M}$ given by the smallest m tried (to obtain a fairly stable estimation).

Finally we have tried a quite local choice for comparison purposes with $m = \frac{1}{6}N$ in the rows labelled 'LOCAL'. We have used 1000 replications for sample sizes $N = 120$ and 256, and 500 for $N = 480$.

Following Bühlmann (1996), for each case we report the bias, standard deviation and the relative mean squared error (normalizing by the optimal or true value) for $\widehat{M}$ and $\widehat{f}_M(0)$. We also give the ratio of the MSE's of the $\widehat{f}_M(0)$ calculated using the optimal choice M^* and the estimated $\widehat{M}$, so a value less than one would indicate a better performance than the one obtained using the (usually unknown) asymptotically optimal value for the bandwidth.

Since we are interested finally in estimating $f(0)$, in our remarks we will concentrate more on the diagnostics for $\widehat{f}_{\widehat{M}}(0)$ than on those for $\widehat{M}$. For Models 1 and 2 and the three sample sizes tried, we observed that almost all the local choices (rows '1' to 'LOCAL') perform better than BB's method ('GLOBAL'), since they adapt to the local properties of the function being estimated. The best procedure varies from case to case, but the simple choices '3' and '4' work uniformly better than the global procedure and seem also to have less variability. Iterating leads to great improvements in some cases, but it does not provide a general advantage. For Model 2 it is possible to observe that the smaller choices of m do not improve from the global procedure, since we are still considering in CVLL_m the peak of the spectrum: here large values of m or just one iteration lead to great improvements in the behaviour of $\widehat{f}_{\widehat{M}}(0)$.

Similar conclusions can be reached for Models 3 to 5, though in Model 4 the procedure '4' (with $m = N^{0.04}$) breaks down, though gives a good initial value for a further iteration. Here we may compare with Bühlmann (1996) results for sample sizes $N = 120$ and 480 (cf. his Table I), although he uses a different class of nonparametric estimates (lag-window or continuously weighted periodogram estimates) with different weight functions. For our Model 3 (Bühlmann's

Model 1) and $N = 120$ all our methods work better (including BB's global choice) both in terms of Rmse and MSE ratio. For $N = 480$, the MSE ratio is still always better (except with the 'LOCAL' choice) and the Rmse is always between the values given by his two proposals.

In the case of Model 4 and $N = 120$ and 480, the methods 'GLOBAL' to '3' worked always better in terms of MSE ratio, but for the smaller sample size gave larger Rmse than Bühlmann's best estimates. The methods with one iteration worked noticeable worse than the ones fully iterated, which only in a few cases outperformed the single-step estimates. For Model 5, Methods '3' and '4' always worked better than any of Bühlmann's alternatives in terms of MSE ratio and also in terms of Rmse for $N = 480$ (like most of the cross-validated choices).

In general, it seems that the asymptotic result for the optimal choice of M for a single frequency is not specially accurate for our periodogram-based estimates, so local cross validation improves even with respect to the knowledge of it (most of the MSE ratio columns have values less than 1, except for Model 4). Also cross validation does not behave never much worse than the iterative plug-in procedure, outperforming it very often.

4.2 Estimation of the whole spectrum

Finally we have tried the local cross-validation for estimation of the optimal bandwidth for all Fourier frequencies λ_j, $j = 0, \ldots, N-1$ for the same models and sample sizes as before. Here the computation costs are much greater, so we have only implemented 200, 100 and 50 simulations for sample sizes 120, 256 and 480 respectively. Given the conclusions of the previous section, we have only tried the Global procedure of BB ($m = 1$) and the three initial choices of $m = 2, 3, 4$ (without iteration), which adapt to the roughness of f at each point. We give in Table VIII the sample mean of the IMSE estimated with the simulations,

$$\frac{1}{N} \sum_{j=0}^{N-1} \left\{ \frac{\widehat{f}_{\widehat{M}}(\lambda_j) - f(\lambda_j)}{f(\lambda_j)} \right\}^2,$$

and its standard deviation.

Almost uniformly the local cross-validation procedures beat the global one, in some situations by a wide margin, and in the worst cases (Models 1 and 5) they perform roughly in the same way. The improvement with respect to the global choice is generally greater the smaller the sample size and, against intuition, in many cases the more local choices also leave to less variable procedures. There are not significative dissimilarities for the three different values of $m > 1$, but $m = 3$ and 4 seem to do slightly better.

Comparing with Bühlmann's Table II, for $N = 120$ and 480 all the local cross-validation IMSE's (and in many cases also BB's global choice) are always better than that of the best plug-in alternative, though they have apparently greater variability, at least in our simulations.

5 Final Remarks

In this paper we have justified a local bandwidth choice procedure for nonparametric spectral estimates and shown its performance in finite sample sizes. We have assumed throughout Gaussianity, but this seems not essential, except perhaps in the proof for the supremum of the periodogram in Lemma 4. We conjecture that this condition can be avoided using Robinson (1991) techniques and assuming summability conditions on higher order cumulants as in Brillinger (1975), except for the second order ones (autocovariances), imposing here only local conditions on the (second order) spectral density.

A multivariate version of the method will be very useful in practical work, but if we want to stress the specific characteristics of each univariate time series it could be better to apply the method to each of them separately or to a fixed linear combination of the series, like in Newey and West (1994).

Further investigation seems necessary in the design of (possibly iterative) algorithms that, linking m and M, reduce the variability inherent to bandwidth choice procedures. Then additional finite sample evidence should be investigated for other models and distributions.

6 Appendix: Proofs and Lemmata

Proof of Lemma 1. An equivalent lemma is evidently valid for more general choices of M, but we are specially interested in this particular case. We can take an $\epsilon > 0$ as small as we want, in such a way that in the interval $I_\nu = [\nu - \epsilon, \nu + \epsilon]$ the conditions of Assumption 2 are satisfied. Then for m big enough we have that $|\nu - \lambda_j| \leq c \cdot m^{-1}$ implies $\lambda_j \in I_\nu$. Therefore when $\nu > 0$ we have that for N big enough, $0 < \lambda_j \sim \nu$, so $(\lambda_j)^{-1} = O(1)$, where $a \sim b$ means $a/b \to 1$ as $N \to \infty$. We study first the bias and the variance.

Bias. Similarly to Theorem 5.6.1 of Brillinger (1975, p.147) and using now Lemma 2 with $\alpha = 1$, we get,

$$\begin{aligned} E[\widehat{f}_M(\lambda_j)] &= \int_{-\pi}^{\pi} K(\lambda) f(\lambda_j - \beta/M) d\beta + O(M/N) \\ &= f(\lambda_j) + \frac{w_2}{2} f^{(2)}(\lambda_j) M^{-2} + O\left(\frac{M}{N} + M^{-3}\right). \end{aligned}$$

The bounded variation condition on K and the derivability of f are used to approximate the discrete average of K and f by and integral with error $O(M/N)$, since by Assumption 2 and for M big enough we are only averaging inside I_ν, thanks to the compact support of K.

Variance. First, it is more convenient to write the spectral estimate using only N frequencies in this way:

$$\widehat{f}_M(\lambda_j) = \frac{\sigma_M^{-1}}{M}\sum_{k=1}^{N-1} K_M(\lambda_k - \lambda_j) I(\lambda_k),$$

where $K_M(\lambda) = \sum_j M\, K(M[\lambda + 2\pi j])$ is the periodic extension of $MK(M\lambda)$. Then we have

$$\begin{aligned} \mathrm{Var}[\widehat{f}_M(\lambda_j)] &= \frac{\sigma_M^{-2}}{M^2}\sum_k K_M(\lambda_k - \lambda_j)^2 \mathrm{Var}[I(\lambda_k)] && (9)\\ &\quad + \frac{\sigma_M^{-2}}{M^2}\sum_k \sum_{i\neq k} K_M(\lambda_k - \lambda_j) K_M(\lambda_i - \lambda_j)\mathrm{Cov}[I(\lambda_k), I(\lambda_i)]. && (10) \end{aligned}$$

Then, from Lemma 3 we get $\mathrm{Var}[I(\lambda_k)] = f(\lambda_k)^2 + O(N^{-1}\log N)$ and, for $k \neq i$,

$$\mathrm{Cov}[I(\lambda_k), I(\lambda_i)] = \begin{cases} f(\lambda_k)^2 + O(N^{-1}\log N) & \text{if} \quad k = N - i \\ O(N^{-1}\log N) & \text{otherwise.} \end{cases}$$

Also we have that $\sigma_M = N/(2\pi M) + O(1)$. Then (9) is

$$\begin{aligned} &\frac{(2\pi)^2}{N^2}\sum_k K_M(\lambda_k - \lambda_j)^2 f(\lambda_k)^2 + O\left(\frac{M}{N^2}\log N\right) \\ &= \frac{2\pi M}{N}\int_{-\pi}^{\pi} MK(M\lambda)^2 f(\lambda_j + \lambda)^2 d\lambda + O\left(\frac{M^2}{N^2}\right) \\ &= \frac{2\pi M}{N} f(\lambda_j)^2 \int_{-\pi}^{\pi} K(\lambda)^2 d\lambda + O\left(\frac{M}{N}\left[\frac{M}{N} + M^{-2}\right]\right). \end{aligned}$$

In (10) we only have to consider the situation where $k = N - i$, since for the other frequencies we have a bound of $O(N^{-1}\log N)$ for the covariance from Lemma 3. Then, if $\nu = 0$ and $\lambda_j = 0$, (10) is similar to (9). In general, if $\nu = 0$ and $|\lambda_j| \leq 2\pi/M$ then the two kernels in (10) overlap in some interval for all M. Taking into account only the frequencies $i = N - k$, for which the leading term of the covariance is also $f(\lambda_k)$ we have that (10) is equal to, using the periodicity of K_M,

$$\begin{aligned} &\frac{\sigma_M^{-2}}{M^2}\sum_k K_M(\lambda_k - \lambda_j) K_M(\lambda_k + \lambda_j)\left[f(\lambda_k)^2 + O(N^{-1}\log N)\right] && (11)\\ &= \frac{2\pi}{N}\int_{-\pi}^{\pi} K_M(\lambda) K_M(\lambda + 2\lambda_j) f(\lambda - \lambda_j)^2 d\lambda + O\left((M/N)^2 + \log N\left[\frac{M}{N^2}\right]\right) \\ &= \delta_M(j) f(\lambda_j)^2 \frac{2\pi M}{N}\int_{-\pi}^{\pi} K(\lambda)^2 d\lambda + O\left(MN^{-1}\left[\frac{M}{N} + M^{-2}\right]\right), \end{aligned}$$

for some $0 < \delta_M(j) \leq 1$. If $|\lambda_j| > 2\pi/M$ then the two kernels in (11) do not overlap at all and the covariance terms do not contribute to the leading term in the variance of $\widehat{f}_M$, and the lemma follows. •

Lemma 2 *Under Assumption 1, if f satisfies a uniform Lipschitz condition of order $0 < \alpha \leq 1$ in an interval around a fixed frequency ν, then for Fourier frequencies such that* $\sup_{\lambda_\ell} |\nu - \lambda_\ell| \leq$

$c \cdot m^{-1}$, $\ell \in \{j, k\}$, *for some positive sequence* m *such that* $1/m + m/N \to 0$, *uniformly in* j *and* k, $(j, k \neq 0)$,

$$E\left[d_x(\lambda_j)\overline{d_x(\lambda_k)}\right] - \delta_{jk}\, 2\pi N f(\lambda_j) = O(N^{1-\alpha}\log N),$$

where $d_x(\lambda_j)$ *is the discrete Fourier transform of the series* X_t,

$$d_x(\lambda_j) = \sum_{t=1}^{N} X_t e^{-i\lambda_j t}.$$

Proof. This lemma is a restatement of, for example, the Lemma in p. 835 of Hannan and Nicholls (1977), assuming only local conditions on f. As in the proof of Lemma 1 we can fix one $\epsilon > 0$ such that, if $I_\nu = [\nu - \epsilon, \nu + \epsilon]$, $\lambda_j, \lambda_k \in I_\nu$ for N big enough. Defining the Dirichlet kernel $\varphi_N(\lambda)$,

$$\varphi_N(\lambda) = \sum_{j=1}^{N} e^{i\lambda j},$$

we have that for $j \neq k$, mod (N),

$$\int_{-\pi}^{\pi} \varphi_N(\lambda_j - \lambda)\varphi_N(\lambda - \lambda_k)d\lambda = 0.$$

Then, if $j \neq k$, mod (N),

$$\mathrm{E}\left[d_x(\lambda_j)\overline{d_x(\lambda_k)}\right] = \int_{-\pi}^{\pi} \varphi_N(\lambda_j - \lambda)\varphi_N(\lambda - \lambda_k)[f(\lambda) - f(\lambda_j)]d\lambda. \tag{12}$$

Now we divide the range of integration in (12) in the following intervals. First,

$$\begin{aligned}\left|\int_{\lambda_j - N^{-1}}^{\lambda_j + N^{-1}} \varphi_N(\lambda_j - \lambda)\varphi_N(\lambda - \lambda_k)[f(\lambda) - f(\lambda_j)]d\lambda\right| &\leq cN \int_{\lambda_j - N^{-1}}^{\lambda_j + N^{-1}} |\lambda - \lambda_j|^{\alpha-1} d\lambda \\ &\leq cN^{1-\alpha},\end{aligned}$$

using $\sup_{\lambda \in I_\nu} |f(\lambda) - f(\lambda_j)| \leq \mathrm{c} \cdot |\lambda - \lambda_j|^{\alpha}$ in the interval considered, and

$$|\varphi_N(\lambda)| \leq \min\left\{2|\lambda|^{-1}, N\right\}.$$

Next,

$$\begin{aligned}&\left|\int_{\lambda_k - N^{-1}}^{\lambda_k + N^{-1}} \varphi_N(\lambda_j - \lambda)\varphi_N(\lambda - \lambda_k)[f(\lambda) - f(\lambda_j)]d\lambda\right| \\ &\leq \mathrm{c} \cdot N^{-1} \sup_{|\lambda - \lambda_k| \leq N^{-1}} |\varphi(\lambda - \lambda_k)| \sup_{|\lambda - \lambda_k| \leq N^{-1}} |\lambda - \lambda_j|^{\alpha - 1} \leq \mathrm{c} \cdot N^{1-\alpha},\end{aligned}$$

since the range of integration was of order N^{-1}. Define the set $I_\nu(k, j)$ as the interval I_ν except the previous two neighbourhoods of radius N^{-1} around λ_k and λ_j. Then

$$\begin{aligned}&\left|\int_{I_\nu(k,j)} \varphi_N(\lambda_j - \lambda)\varphi_N(\lambda - \lambda_k)[f(\lambda) - f(\lambda_j)]d\lambda\right| \\ &\leq \mathrm{c} \sup_{I_\nu(k,j)} |\lambda - \lambda_j|^{\alpha-1} \int_{-\pi}^{\pi} |\varphi_N(\lambda - \lambda_k)| d\lambda \\ &\leq \mathrm{c} \cdot N^{1-\alpha}\log N,\end{aligned}$$

using $\int_{-\pi}^{\pi}|\varphi_N(\lambda)|d\lambda \leq c \cdot \log N$. Finally in the complementary set of I_ν,

$$\left|\int_{I_\nu^c}\varphi_N(\lambda_j-\lambda)\varphi_N(\lambda-\lambda_k)[f(\lambda)-f(\lambda_j)]d\lambda\right|$$
$$\leq c\sup_{I_\nu^c}|\varphi_N(\lambda_j-\lambda)\varphi_N(\lambda-\lambda_k)|\left[f(\lambda_j)+\int_{-\pi}^{\pi}f(\lambda)d\lambda\right]\leq c,$$

and the lemma follows in the case $j \neq k$ because any of the bounds depends on j or k. If $j = k$ then we can use the same methods as before together with

$$\int_{-\pi}^{\pi}|\varphi_N(\lambda)|^2 d\lambda = 2\pi N,$$

to get the desired result. •

Lemma 3 *Under Assumption 1, if f satisfies a uniform Lipschitz condition of order $0 < \alpha \leq 1$, in an interval around a fixed frequency ν and if $\sup_{\lambda_{j_r}}|\nu-\lambda_{j_r}| \leq c \cdot m^{-1}$, $r = 1,\ldots,q$, for some positive sequence m such that $1/m + m/N \to 0$, then, uniformly in $j_r \neq 0$, with $j_r \neq j_{r'}$, $r \neq r'$,*

$$E\left[\prod_{r=1}^{q} I(\lambda_{j_r})^{p_r}\right] = \prod_{r=1}^{q} p_r! f(\lambda_{j_r})^{p_r} + O(N^{-\alpha}\log N), \tag{13}$$

and

$$E\left[\prod_{r=1}^{q}\left(\frac{I(\lambda_{j_r})-f(\lambda_{j_r})}{f(\lambda_{j_r})}\right)\right] = O(N^{-\alpha}\log N). \tag{14}$$

Proof. The proof it is immediate in the light of the Proposition in page 31 of BB and our Lemma 2, as by the Gaussianity of X_t only cumulants of order two of the discrete Fourier transform of X_t have to be considered. Here the bound in (14) is only $O(N^{-\alpha}\log N)$ and not this bound to the power of q as in BB. The problem with their proof is the following. At the beginning of their page 33, for $k \in \nu_2$ in their notation, $\mathrm{cum}\{d_x(\lambda_{k1}), d_x(\lambda_{k2})\}N^{-1} = O(1)$ at most, because we can have $\lambda_{k1} = \lambda_{k2}$ for all elements in one of the possible partitions. Then, the second bound in the third full paragraph formula of the same page is only $O(1)$ and the first one is $O(N^{-\alpha}\log N)$ (actually $O(N^{-1})$ under their conditions), since we have $\#\nu_1 \leq 1$. •

We give now some lemmas needed for the proof of Proposition 1.

Lemma 4 *Under Assumptions 1, if f satisfies a uniform Lipschitz condition of order $0 < \alpha \leq 1$, in an interval around a fixed frequency ν, $I_\nu = [\nu-\epsilon, \nu+\epsilon]$ for some $\epsilon > 0$, then for frequencies $\lambda_j = 2\pi j/N$, $j \neq 0$ such that $\sup_{\lambda_j}|\nu-\lambda_j| \leq c \cdot m^{-1}$, for some positive sequence m such that $1/m + m/N \to 0$, uniformly in $j \neq 0$,*

$$\overline{\lim_{N\to\infty}}\sup_{\lambda_j} I(\lambda_j) \leq 2\log N \sup_{\lambda\in I_\nu} f(\lambda) \quad \textit{w.p.1.}$$

Proof. We can proceed as in the proof of Theorems 4.5.1 and 5.3.2 of Brillinger (1975), taking the mean of X_t as zero, since we do not include the zero frequency. In our case, since X_t is a Gaussian series and $j \neq 0$, all the cumulants of order bigger than two are zero. From Lemma 2 we can obtain, uniformly in j, for m big enough,

$$\sigma_N \equiv \mathrm{Var}[\mathrm{Re}\, d_x(\lambda_j)] = \frac{N}{2} 2\pi f(\lambda_j) + O(N^{1-\alpha} \log N).$$

Then, for $\lambda_j \in I_\nu$ and any θ and one ϵ as small as we want, from Gaussianity, as $N \to \infty$,

$$\mathrm{E}\left[\exp\left\{\theta \mathrm{Re}\, d_x(\lambda_j)\right\}\right] \leq \exp\left\{\theta^2 2\pi N f(\lambda_j)(1+\epsilon)/4\right\}.$$

Next,

$$\begin{aligned} \mathrm{E}\exp\left\{\theta \sup_{\lambda_j} |\mathrm{Re}\, d_x(\lambda_j)|\right\} &\leq \sum_{\lambda_j \in I_\nu} \mathrm{E}\exp\left\{\theta\, |\mathrm{Re}\, d_x(\lambda_j)|\right\} \\ &\leq \sum_{\lambda_j \in I_\nu} \exp\left\{\theta^2 2\pi N f(\lambda_j)(1+\epsilon)/4\right\} \\ &\leq 2\exp\left\{\log N + \theta^2 2\pi N \sup_{\lambda \in I_\nu} f(\lambda)(1+\epsilon)/4\right\}. \end{aligned}$$

Now define, for $\delta > 0$

$$a^2 = 2\pi(1+\epsilon)(2+\delta) N \log N \sup_{\lambda \in I_\nu} f(\lambda).$$

Then

$$\mathrm{Prob}\left\{\sup_{\lambda_j} |\mathrm{Re}\, d_x(\lambda_j)| \geq a\right\} \leq \exp\{-\theta a\} 2\exp\left\{\log N + \theta^2 2\pi N \sup_{\lambda \in I_\nu} f(\lambda)(1+\epsilon)/2\right\}.$$

Taking

$$\theta = a\left[2\pi N(1+\epsilon) \sup_{\lambda \in I_\nu} f(\lambda)\right]^{-1},$$

this is less or equal than

$$2\exp\left\{-a^2\left[2\pi N \sup_{\lambda \in I_\nu} f(\lambda)(1+\epsilon)\right]^{-1}\right\} \exp\{\log N\} \leq \mathrm{c} \cdot N^{-1-\delta}.$$

Using this last line and the Borel-Cantelli Lemma, as ϵ and δ were arbitrary, we obtain that

$$\overline{\lim_{N\to\infty}} \sup_{\lambda_j} |\mathrm{Re}\, d_x(\lambda_j)| / [2\pi N \log N]^{1/2} \leq \left[\sup_{\lambda \in I_\nu} f(\lambda)\right]^{1/2} \quad \text{w.p.1.}$$

A similar result is possible for the imaginary part of d_x and then the lemma follows from

$$|d_x(\lambda_j)| \leq |\mathrm{Re}\, d_x(\lambda_j)| + |\mathrm{Im}\, d_x(\lambda_j)|$$

and $I(\lambda_j) = (2\pi N)^{-1} |d_x(\lambda_j)|^2$. •

Lemma 5 *Under Assumptions 1, 2, 3, 4, for frequencies* $\lambda_j = 2\pi j/N$ *such that* $\sup_{\lambda_j} |\nu - \lambda_j| \leq c \cdot m^{-1}$, *for some positive sequence* m *such that* $1/m + m/N \to 0$, *uniformly in* j,

$$\sup_{\lambda_j} \left| \frac{\widehat{f}_M(\lambda_j) - f(\lambda_j)}{f(\lambda_j)} \right| = O_P\left(N^{-1}M^2 + N^{\frac{1-p}{2p}} M + N^{-1} \log N + M^{-1} \right) = o_P(1).$$

Proof. Define the weighted autocovariance spectral estimate corresponding to the continuous average in $\widehat{f}_M$, when the mean of X_t is known, (and assumed to be 0 without loss of generality),

$$\widehat{f}_M^C(\lambda_j) = \int_{-\pi}^{\pi} K_M(\lambda) I(\lambda_j + \lambda) d\lambda = \frac{1}{2\pi} \sum_{r=1-N}^{N-1} w\left(\frac{r}{M}\right) \widehat{\gamma}(r) \cos r\lambda_j,$$

where $K_M(\cdot) = MK(M\cdot)$ periodically extended and

$$\widehat{\gamma}(k) = N^{-1} \sum_{1 \leq t, t+k \leq N} X_t X_{t+k}.$$

This estimate is unfeasible if the mean of the series is unknown, but we only need its definition for the proofs. Now we have, proceeding as in the proof of Theorem 2.1 of Robinson (1991),

$$\begin{aligned}
\sup_{\lambda_j} \left| \widehat{f}_M(\lambda_j) - f(\lambda_j) \right| &\leq \sup_{\lambda_j} \left| \widehat{f}_M(\lambda_j) - \widehat{f}_M^C(\lambda_j) \right| && (15) \\
&\quad + \sup_{\lambda_j} \left| \widehat{f}_M^C(\lambda_j) - E\left[\widehat{f}_M^C(\lambda_j)\right] \right| && (16) \\
&\quad + \sup_{\lambda_j} \left| E\left[\widehat{f}_M^C(\lambda_j)\right] - f(\lambda_j) \right|. && (17)
\end{aligned}$$

Now (15) is less or equal than (see Robinson, 1991, p. 1353),

$$\begin{aligned}
(2\pi)^{-1} \sum_{1-N}^{N-1} \left| w\left(\frac{r}{M}\right) \right| |\widehat{\gamma}(N-r)| &= O_P\left(N^{-1} \sum_{1-N}^{N-1} \left| w\left(\frac{r}{M}\right) \right| |r| \right) \\
&= O_P\left(N^{-1} M^{\alpha} \sum_{1}^{N} |r|^{1-\alpha} + N^{-1} M^2 \log N \right) \\
&= O_P\left(N^{1-\alpha} M^{\alpha} + N^{-1} M^2 \log N \right) = o_P(1),
\end{aligned}$$

using Assumption 4 ($\alpha > \frac{5}{4}$) and the fact that $\widehat{\gamma}(N-r)$ is a sum of r terms whose mean exists and is uniformly bounded. Next (16) is not bigger than

$$\begin{aligned}
(2\pi)^{-1} \sum_{1-N}^{N-1} \left| w\left(\frac{r}{M}\right) \right| |\widehat{\gamma}(r) - E[\widehat{\gamma}(r)]| &= O_P\left(\sum_{1-N}^{N-1} \left| w\left(\frac{r}{M}\right) \right| N^{\frac{1-p}{2p}} \right) \\
&= O_P\left(N^{\frac{1-p}{2p}} M \right) = o_P(1),
\end{aligned}$$

because Assumptions 2 and 4, and Lemma 7 below. Finally (17) is bounded by

$$\begin{aligned}
&\sup_{\lambda_j} \left| \int_{-\pi}^{\pi} K_M(\lambda_j - \omega) \left\{ E[I(\omega) - f(\omega)] \right\} d\omega \right| && (18) \\
&+ \sup_{\lambda_j} \left| \int_{\mathcal{R}} K(\omega) \left\{ f(\lambda_j - \omega/M) - f(\omega) \right\} d\omega \right|. && (19)
\end{aligned}$$

Denote by Φ_N Fejér Kernel $\Phi_N(\lambda) = (2\pi N)^{-1}|\varphi_N(\lambda)|^2$. Similarly to Lemma 2, we have that in (18) ω lies in the interior of I_ν as $M \to \infty$ due to the compact support of K, and for fixed $\delta > 0$ small enough,

$$\begin{aligned}
\sup_{\omega\in I_\nu} |E[I(\omega)] - f(\omega)| &\leq \sup_{\omega\in I_\nu} \left| \int_{-\pi}^{\pi} \Phi_N(\alpha-\omega)[f(\alpha) - f(\omega)]d\alpha \right| \\
&\leq \sup_{\omega\in I_\nu} |f'(\omega)| \int_{|\omega-\alpha|\leq\delta} |\Phi_N(\alpha-\omega)||\alpha-\omega|d\alpha \\
&\quad + \sup_{\omega\in I_\nu} \int_{|\omega-\alpha|>\delta} |\Phi_N(\alpha-\omega)|[f(\alpha) + f(\omega)]d\alpha \\
&= O(N^{-1}\log N) + O(N^{-1}) \\
&= O(N^{-1}\log N),
\end{aligned}$$

uniformly in $\omega \in I_\nu$, so (18) is $O(N^{-1}\log N)$, since $\int |K_M(\alpha)|d\alpha < \infty$. Next, as $M \to \infty$, (19) is bounded by (denoting by λ^* a value between λ_j and $\lambda_j - \omega/M$),

$$\sup_{\lambda_j} \int K(\omega)\left[f(\lambda_j - \omega/M) - f(\lambda_j)\right] d\omega \leq \sup_{\lambda_j} \int |K(\omega)||f'(\lambda^*)| \left|\frac{\omega}{M}\right| d\omega = O(M^{-1}),$$

using the compact support of K and that of f' is bounded in I_ν. •

Lemma 6 *Under the Assumptions of Lemma 5, uniformly in j,*

$$\sup_{\lambda_j} \left| \frac{\widehat{f}^{\mathbf{j}}_M(\lambda_j) - f(\lambda_j)}{f(\lambda_j)} \right| = O_P\left(N^{-1}M^2 + N^{\frac{1-p}{2p}}M + N^{-1}\log N + M^{-1}\right).$$

Proof. The proof is exactly the same as that of Lemma 4 of BB, using now our Lemma 5. •

Lemma 7 *Under Assumptions 1 and 2, uniformly in r, $p > 1$,*

$$Var[\widehat{\gamma}(r)] = O\left(N^{\frac{1-p}{p}}\right).$$

where $\widehat{\gamma}(r)$ is the (biased) estimate of the lag-r autocovariance $\gamma(r)$ when the expectation of X_t is known.

$$\widehat{\gamma}(r) = \frac{1}{N} \sum_{1\leq t,t+r\leq N} (X_t - E[X_1])(X_{t+r} - E[X_1]).$$

Proof. From e.g. Anderson (1971, p. 452), denoting as before the Fejér kernel by Φ_N,

$$N\,\mathrm{Var}\left[\widehat{\gamma}(r)\right] = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \Phi_N(\alpha-\omega)(1 + e^{-i(\alpha+\omega)r})f(\alpha)f(\omega)d\alpha d\omega,$$

and now the result follows applying Hölder inequality twice, with $|\Phi_N(\omega)| = O(N)$ uniformly in ω, $|1 + e^{-i(\alpha+\omega)r}| \leq 2$ uniformly in r and $\int_{-\pi}^{\pi} f^p < \infty$ by Assumption 2. •

Proof of Proposition 1. From the proof of Theorem 3.1 in BB the proposition will follow, using their definitions, if we show

$$\begin{aligned}
N^{-1}T_i &= o_P(\mathrm{IMSE}_m) \qquad i = 1,2 \\
N^{-1}T_3 &= \mathrm{IMSE}_m + o_P(\mathrm{IMSE}_m).
\end{aligned}$$

First we have, denoting now $\sigma_j = \sigma_{j,M}$, from the last steps in the proofs of BB,

$$E[T_1] = 2\pi \sum_{j=1}^{N} W_m(\lambda_j - \nu)\sigma_j^{-1} {\sum_k}' K(M\lambda_k) O(N^{-1}) = O(1),$$

and, denoting as $\mathrm{IMSE}'_m(\nu, M)$ the IMSE_m calculated from the modified spectral estimate (1),

$$\begin{aligned}
E[T_1^2] &= N\,\mathrm{MISE}'_m + 2\pi \sum_j W_m(\lambda_j - \nu)^2 \sigma_j^{-2} {\sum_k}'{\sum_n}' K(M\lambda_k)K(M\lambda_n)O(N^{-1}) \\
&\quad + 2\pi \sum_j \sum_{j \neq i} W_m(\lambda_j - \nu) W_m(\lambda_i - \nu) \sigma_j^{-1}\sigma_i^{-1} {\sum_k}'{\sum_n}' K(M\lambda_k)K(M\lambda_n)O(N^{-2}) \\
&\quad + 2\pi \sum_j W_m(\lambda_j - \nu)^2 \sigma_j^{-2} {\sum_k}'{\sum_n}' K(M\lambda_k)^2 \\
&= N\,\mathrm{IMSE}'_m + O(m) + O(m) + O(m\,N\,\mathrm{IMSE}_m) \\
&= O(m\,N\,\mathrm{IMSE}_m),
\end{aligned}$$

since $\sup_{\lambda,m} |W_m(\lambda)| = O(m)$. Then using $\mathrm{IMSE}_m = O(M/N)$ we can obtain

$$T_1 = O_P(\mathrm{IMSE}_m\,[m/M]^{1/2}) = o_P(\mathrm{IMSE}_m),$$

because $m/M \to 0$. Now, in a similar fashion,

$$E[T_2] = 2\pi \sum_{j=1}^{N} W_m(\lambda_j - \nu)\sigma_j^{-2} {\sum_k}'{\sum_n}' K(M\lambda_k)K(M\lambda_n)O(N^{-1}) = O(1),$$

and as before

$$E[T_2] = O(m\,N\,\mathrm{IMSE}_m^2).$$

(Note that in BB's expression they have N^{-1} instead of N in the correspondent formula, although in the statements in the main part of their paper they give the right bounds). Then $N^{-1}T_2 = O_P(\mathrm{IMSE}_m\,[m/N]^{1/2}) = o_P(\mathrm{IMSE}_m)$. Next,

$$\begin{aligned}
E[T_3] &= N\,\mathrm{IMSE}_m + O([N/M]^{-1}\,N\,\mathrm{IMSE}_m) + O([N/M]^{-1}) + O([N/M]^{-1}\,M) \\
&= N\,\mathrm{IMSE}_m + O([N/M]^{-1}\,N\,\mathrm{IMSE}_m),
\end{aligned}$$

and reasoning in the same way as before,

$$\mathrm{Var}[T_3] = O(m\,N\,\mathrm{IMSE}_m).$$

Then $N^{-1}T_3 = \mathrm{IMSE}_m + O_P(\mathrm{IMSE}_m[m/M]^{1/2}) = \mathrm{IMSE}_m + o_P(\mathrm{IMSE}_m)$. The proof for the remainder term in BB's expression (3.2) continues the same here, using now our Lemmas 2, 4 and 5 instead of their references, since the bound for the third term in the expansion still holds for the modified (local) cross-validation. •

References

[1] ANDERSON, T.W. (1971) The Statistical Analysis of Time Series. Wiley, New York.

[2] ANDREWS, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 60, 953-966.

[3] BELTRAO, K.I. and BLOOMFIELD, P. (1987) Determining the bandwidth of a kernel spectrum estimate. *Journal of Time Series Analysis* 8, 21-38.

[4] BRILLINGER, D.R. (1975) *Time Series: Data Analysis and Theory.* Holt, Rinehart and Winston, New York.

[5] BROCKMANN, M., GASSER, T. and HERRMANN, E. (1993) Locally adaptive bandwidth choice for kernel regression estimators. *Journal of the American Statistical Association* 88, 1302-1309.

[6] BÜHLMANN, P. (1996) Locally adaptive lag-window spectral estimation. *Journal of Time Series Analysis* 17, 247-270.

[7] Den HAAN, W.J. and LEVIN, A. (1996) A practitioners guide to robust covariance matrix estimation. Discussion Paper, Department of Economics, University of California, San Diego.

[8] HANNAN, E.J. and NICHOLLS, D.F. (1977) The estimation of the prediction error variance. *Journal of the American Statistical Association* 72, 834-840.

[9] HURVICH, C.M. (1985) Data driven choice of a spectrum estimate: extending the applicability of cross-validation methods. *Journal of the American Statistical Association* 80, 933-940.

[10] HERRMANN, E. (1997) Local bandwidth choice in kernel regression estimation. *Journal of Computational and Graphical Statistics* 6, 35-54.

[11] HURVICH, C.M. and BELTRAO, K.I. (1994) Automatic semiparametric estimation of the memory parameter of a long-memory time series. *Journal of Time Series Analysis* 15, 285-302.

[12] LEPSKII, O.V. and SPOKOINY, V.G. (1995) Local adaptivity to inhomogeneous smoothness.1. Resolution Level. Preprint.

[13] NEWEY, W.K. and WEST, K.D. (1994) Automatic lag selection in covariance matrix estimation *Review of Economic Studies* 61, 631-653.

[14] ROBINSON, P.M. (1991) Automatic frequency domain inference on semiparametric and nonparametric models. *Econometrica* 59, 1329-1363.

[15] WAHBA, G. (1980) Automatic smoothing of the log periodogram. *Journal of the American Statistical Association* 75, 122-132.

Table I M minimizing $\mathrm{IMSE}_m(0, M)$

		MODEL 1		MODEL 2		MODEL 3		MODEL 4		MODEL 5	
$M^\star(0)$:		7.158		7.131		20.340		11.076		11.969	
m	band	M	bias	M	bias	M	bias	M	bias	M	bias
1	129	7.00	-0.16	24.00	16.87	14.50	-5.84	12.00	0.92	12.25	0.28
1.02	125	7.00	-0.16	24.50	17.37	15.50	-4.84	13.00	1.92	12.25	0.28
1.06	121	7.00	-0.16	24.50	17.37	15.50	-4.84	13.00	1.92	12.25	0.28
1.09	117	7.00	-0.16	24.50	17.37	15.50	-4.84	13.00	1.92	12.25	0.28
1.13	113	7.00	-0.16	25.00	17.87	15.50	-4.84	13.00	1.92	12.25	0.28
1.17	109	7.00	-0.16	25.00	17.87	15.50	-4.84	13.00	1.92	12.25	0.28
1.21	105	7.00	-0.16	27.50	20.37	15.50	-4.84	13.50	2.42	12.00	0.03
1.26	101	7.00	-0.16	27.50	20.37	15.50	-4.84	13.50	2.42	12.00	0.03
1.31	97	6.50	-0.66	27.50	20.37	15.50	-4.84	13.50	2.42	12.00	0.03
1.37	93	6.50	-0.66	27.50	20.37	15.50	-4.84	13.50	2.42	12.00	0.03
1.43	89	6.50	-0.66	27.50	20.37	15.50	-4.84	13.50	2.42	12.00	0.03
1.50	85	1.50	-5.66	28.00	20.87	15.50	-4.84	13.50	2.42	12.00	0.03
1.58	81	1.50	-5.66	28.00	20.87	15.50	-4.84	13.50	2.42	12.00	0.03
1.66	77	1.50	-5.66	28.00	20.87	15.50	-4.84	13.50	2.42	12.00	0.03
1.76	73	4.00	-3.16	28.50	21.37	15.50	-4.84	13.50	2.42	12.25	0.28
1.85	69	4.50	-2.66	28.50	21.37	16.50	-3.84	14.00	2.92	12.25	0.28
1.97	65	5.00	-2.16	28.50	21.37	16.50	-3.84	14.00	2.92	12.25	0.28
2.09	61	6.00	-1.16	24.00	16.87	17.00	-3.34	14.00	2.92	11.25	-0.72
2.24	57	6.00	-1.16	24.00	16.87	17.00	-3.34	14.00	2.92	11.25	-0.72
2.41	53	6.50	-0.66	23.00	15.87	17.00	-3.34	14.00	2.92	11.25	-0.72
2.61	49	6.50	-0.66	23.00	15.87	17.00	-3.34	14.00	2.92	11.00	-0.97
2.84	45	6.50	-0.66	22.50	15.37	17.00	-3.34	14.00	2.92	9.50	-2.47
3.12	41	6.00	-1.16	15.00	7.87	1.50	-18.84	12.50	1.42	3.00	-8.97
3.45	37	6.00	-1.16	13.50	6.37	1.50	-18.84	12.00	0.92	3.00	-8.97
3.87	33	5.50	-1.66	12.00	4.87	1.50	-18.84	11.00	-0.08	3.00	-8.97
4.41	29	5.00	-2.16	10.00	2.87	1.50	-18.84	11.00	-0.08	2.75	-9.22
5.12	25	4.00	-3.16	9.00	1.87	1.50	-18.84	10.50	-0.58	2.50	-9.22
6.09	21	3.00	-4.16	8.50	1.37	1.50	-18.84	1.50	-9.58	2.50	-9.22
7.52	17	3.00	-4.16	8.00	0.87	1.50	-18.84	1.50	-9.58	2.75	-9.22
9.84	13	2.50	-4.66	7.50	0.37	1.50	-18.84	1.50	-9.58	2.75	-9.22
14.22	9	2.00	-5.16	7.00	-0.13	2.00	-18.34	1.50	-9.58	3.00	-8.97
25.60	5	2.00	-5.16	6.50	-0.63	3.50	-16.84	1.50	-9.58	3.00	-8.97
128.0	1	2.00	-5.16	6.50	-0.63	11.00	-9.34	1.50	-9.58	3.00	-8.97

Table II M minimizing $\mathrm{CVLL}_m(0, M)$

		MODEL 1			MODEL 2			MODEL 3			MODEL 4			MODEL 5		
M*(0):			7.158			7.131			20.340			11.076			11.969	
m	band	bias	sd	mse	bias	sd	mse	bias	sd	mse	bias	sd	mse	bias	sd	mse
1.11	115	3.42	4.20	29.35	9.54	4.54	111.52	-10.98	4.08	137.08	-0.49	3.72	14.10	-0.13	3.70	13.74
1.17	109	3.47	4.22	29.82	9.76	4.59	116.36	-10.85	4.24	135.62	-0.45	3.68	13.76	-0.04	3.85	14.81
1.24	103	3.49	4.19	29.74	9.96	4.57	120.13	-10.74	4.32	134.08	-0.28	3.88	15.11	-0.01	3.94	15.54
1.32	97	3.52	4.27	30.64	10.11	4.67	124.04	-10.74	4.27	133.53	-0.13	4.00	15.98	0.05	3.85	14.84
1.41	91	3.68	4.38	32.69	10.54	4.85	134.69	-10.55	4.37	130.44	0.03	3.99	15.94	-0.06	3.89	15.15
1.51	85	3.80	4.36	33.48	11.01	5.04	146.66	-10.41	4.46	128.19	0.12	4.11	16.90	-0.30	4.05	16.52
1.62	79	3.83	4.36	33.69	11.55	5.52	163.94	-10.34	4.27	125.14	0.32	4.12	17.11	-0.18	4.15	17.22
1.75	73	3.91	4.45	35.09	11.88	5.56	172.01	-10.22	4.38	123.63	0.47	4.09	16.98	0.00	4.44	19.72
1.91	67	3.92	4.63	36.79	12.33	5.67	184.10	-10.05	4.54	121.58	0.74	4.45	20.39	-0.01	4.55	20.71
2.10	61	3.61	5.05	38.56	12.80	6.12	201.33	-9.90	4.52	118.52	0.95	4.53	21.42	0.02	4.76	22.70
2.33	55	3.08	5.62	41.13	13.08	6.55	213.86	-9.73	4.62	116.11	1.19	4.87	25.15	0.11	5.10	26.06
2.61	49	2.80	6.18	46.00	10.83	5.93	152.52	-9.41	4.94	112.97	1.28	4.89	25.49	0.12	5.37	28.83
2.98	43	2.85	6.13	45.74	9.95	5.31	127.17	-9.17	5.08	109.82	1.52	5.09	28.21	0.42	5.68	32.46
3.46	37	2.98	6.23	47.65	6.26	4.64	60.72	-8.66	5.53	105.52	1.91	5.87	38.12	0.81	6.18	38.88
4.13	31	2.92	6.27	47.85	4.32	4.48	38.69	-8.58	5.57	104.62	2.12	5.68	36.79	-0.27	6.60	43.68
5.12	25	3.02	6.83	55.78	2.97	5.16	35.43	-7.84	6.10	98.70	1.86	6.20	41.91	-1.45	7.11	52.70
6.74	19	-3.58	3.44	24.66	0.87	4.12	17.70	-8.26	6.05	104.87	-4.50	5.25	47.76	-8.14	2.90	74.61
9.85	13	-2.18	3.82	19.32	-0.14	4.01	16.08	-6.64	7.10	94.61	-2.95	6.04	45.25	-7.54	3.04	66.15
18.29	7	0.38	4.65	21.76	0.06	5.51	30.40	-6.85	8.29	115.77	-1.34	6.21	40.33	-4.58	5.31	49.24
128.0	1	9.15	10.39	191.59	4.00	9.41	104.49	-3.97	9.78	111.51	5.54	10.11	132.96	4.24	10.08	119.50

Table III $\widehat{M}$, $\hat{f}_{\widehat{M}}(0)$ MODEL 1. AR(3) α = [0.60, -0.60, 0.30]

$N = 120$	M*(0): 6.1516			f(0): 0.3248				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	0.3584	2.1217	0.1224	-0.0177	0.1263	0.1542	(0.0355)	1.2357
1	0.0356	2.3079	0.1408	-0.0194	0.1228	0.1465	(0.0344)	1.1746
2	-0.2734	2.3369	0.1463	-0.0204	0.1191	0.1385	(0.0335)	1.1100
3	-2.8218	1.8509	0.3009	-0.0334	0.0872	0.0826	(0.0253)	0.6618
4	-2.8278	1.8884	0.3055	-0.0406	0.0930	0.0977	(0.0264)	0.7827
5	-2.8209	1.7909	0.2950	-0.0397	0.0907	0.0929	(0.0259)	0.7444
6	-2.8381	1.8012	0.2986	-0.0397	0.0905	0.0926	(0.0258)	0.7424
7	-2.8488	1.7770	0.2979	-0.0398	0.0904	0.0926	(0.0258)	0.7419
8	-1.2422	2.4017	0.1932	-0.0183	0.1171	0.1332	(0.0327)	1.0678
9	-1.1877	2.4329	0.1937	-0.0184	0.1223	0.1449	(0.0335)	1.1618
ITER.1	-1.8846	2.4752	0.2558	-0.0284	0.1135	0.1297	(0.0302)	1.0689
ITER.2	-1.8603	2.4889	0.2551	-0.0275	0.1134	0.1291	(0.0311)	1.0638
ITER.3	-1.0683	2.3353	0.1743	-0.0285	0.1148	0.1326	(0.0312)	1.0927
LOCAL	-2.7154	2.1030	0.3117	-0.0472	0.0921	0.1016	(0.0276)	0.8375
$N = 256$	M*(0): 7.1582			f(0): 0.3248				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	1.2240	2.3207	0.1343	-0.0156	0.0918	0.0821	(0.0269)	1.2578
1	1.0328	2.5056	0.1433	-0.0154	0.0919	0.0822	(0.0269)	1.2593
2	0.2098	2.5265	0.1254	-0.0165	0.0877	0.0754	(0.0257)	1.1553
3	-3.6296	2.1438	0.3468	-0.0308	0.0586	0.0416	(0.0174)	0.6371
4	-4.0582	1.8538	0.3885	-0.0383	0.0596	0.0476	(0.0177)	0.7288
5	-3.9735	1.7898	0.3707	-0.0387	0.0598	0.0481	(0.0177)	0.7363
6	-3.9495	1.7956	0.3673	-0.0386	0.0599	0.0481	(0.0177)	0.7368
7	-3.9461	1.7606	0.3644	-0.0377	0.0604	0.0481	(0.0179)	0.7373
8	-1.2117	2.7952	0.1811	-0.0135	0.0865	0.0727	(0.0253)	1.1136
9	-1.1019	2.9139	0.1894	-0.0123	0.0872	0.0736	(0.0255)	1.1266
ITER.1	-1.9004	3.2307	0.2742	-0.0121	0.0864	0.0722	(0.0253)	1.0665
ITER.2	-1.8546	3.2277	0.2704	-0.0120	0.0867	0.0726	(0.0235)	1.0725
ITER.3	-0.4209	2.8948	0.1670	-0.0104	0.0876	0.0738	(0.0245)	1.0902
LOCAL	-3.7812	1.9504	0.3533	-0.0327	0.0607	0.0450	(0.0186)	0.6651
$N = 480$	M*(0): 8.1171			f(0): 0.3248				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	1.7736	2.5416	0.1458	-0.0111	0.0736	0.0525	(0.0308)	1.3254
1	1.6460	2.5955	0.1434	-0.0107	0.0738	0.0527	(0.0309)	1.3304
2	0.5319	2.7710	0.1208	-0.0121	0.0699	0.0478	(0.0293)	1.2052
3	-4.2811	2.4304	0.3678	-0.0269	0.0412	0.0230	(0.0174)	0.5795
4	-5.2523	1.6165	0.4584	-0.0354	0.0410	0.0279	(0.0173)	0.7031
5	-5.1230	1.6590	0.4401	-0.0365	0.0409	0.0285	(0.0173)	0.7204
6	-5.0873	1.7477	0.4392	-0.0368	0.0406	0.0285	(0.0171)	0.7188
7	-4.9684	1.7539	0.4213	-0.0362	0.0420	0.0291	(0.0177)	0.7340
8	-1.2159	3.3937	0.1972	-0.0118	0.0683	0.0456	(0.0286)	1.1495
9	-0.8354	3.3192	0.1778	-0.0071	0.0683	0.0446	(0.0285)	1.1266
ITER.1	-1.8274	3.7082	0.2594	-0.0103	0.0681	0.0449	(0.0282)	1.1118
ITER.2	-1.7917	3.7178	0.2585	-0.0101	0.0680	0.0447	(0.0284)	1.1078
ITER.3	-0.0604	3.1202	0.1478	-0.0106	0.0690	0.0462	(0.0292)	1.1449
LOCAL	-4.8619	2.1285	0.4275	-0.0289	0.0429	0.0254	(0.0182)	0.6285

Table IV $\widehat{M}$, $\hat{f}_{\widehat{M}}(0)$ MODEL 2. AR(2) α = [0.60, -0.90]

$N = 120$	M*(0): 6.1285			f(0): 0.0942				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	4.9919	1.3243	0.7102	0.0144	0.0580	0.4027	(0.2042)	1.4147
1	5.0619	1.3069	0.7277	0.0145	0.0583	0.4075	(0.2054)	1.4315
2	5.0557	1.3236	0.7272	0.0146	0.0582	0.4063	(0.2050)	1.4275
3	4.0417	1.7333	0.5149	0.0157	0.0563	0.3858	(0.1986)	1.3555
4	0.8858	2.2370	0.1541	0.0179	0.0517	0.3378	(0.1825)	1.1867
5	0.5057	2.4880	0.1716	0.0148	0.0481	0.2856	(0.1705)	1.0035
6	0.5049	2.5140	0.1751	0.0148	0.0483	0.2878	(0.1712)	1.0111
7	0.4994	2.5041	0.1736	0.0149	0.0483	0.2879	(0.1712)	1.0114
8	0.5482	2.4251	0.1646	0.0153	0.0483	0.2898	(0.1713)	1.0182
9	1.8352	1.9411	0.1900	0.0163	0.0501	0.3123	(0.1774)	1.0971
ITER.1	1.5649	2.4872	0.2299	0.0198	0.0544	0.3779	(0.0369)	1.3026
ITER.2	1.5576	2.5005	0.2311	0.0197	0.0543	0.3767	(0.0365)	1.2985
ITER.3	2.3805	2.0955	0.2678	0.0171	0.0541	0.3624	(0.0364)	1.2493
LOCAL	0.1151	2.7552	0.2025	0.0203	0.0532	0.3656	(0.0345)	1.2603
$N = 256$	M*(0): 7.1313			f(0): 0.0942				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	8.8523	2.2880	1.6439	0.0061	0.0395	0.1802	(0.1403)	1.7627
1	9.0051	2.2794	1.6967	0.0061	0.0397	0.1822	(0.1411)	1.7823
2	8.8244	2.3557	1.6403	0.0060	0.0397	0.1820	(0.1410)	1.7795
3	5.5188	2.4520	0.7171	0.0068	0.0362	0.1529	(0.1285)	1.4949
4	0.3458	2.9035	0.1681	0.0079	0.0311	0.1157	(0.1104)	1.1317
5	0.0524	3.1556	0.1959	0.0080	0.0309	0.1150	(0.1100)	1.1246
6	0.0325	3.1302	0.1927	0.0079	0.0309	0.1147	(0.1099)	1.1221
7	0.0752	3.1215	0.1917	0.0077	0.0309	0.1145	(0.1099)	1.1193
8	0.2211	2.9473	0.1718	0.0075	0.0305	0.1112	(0.1084)	1.0872
9	2.3715	2.7299	0.2571	0.0086	0.0336	0.1355	(0.1193)	1.3250
ITER.1	1.6299	3.4685	0.2888	0.0110	0.0356	0.1567	(0.0123)	1.3659
ITER.2	1.6644	3.4757	0.2920	0.0109	0.0356	0.1564	(0.0122)	1.3631
ITER.3	2.7699	3.1770	0.3493	0.0094	0.0351	0.1492	(0.0124)	1.3004
ITER.4	-0.3036	3.2764	0.2129	0.0151	0.0361	0.1731	(0.0128)	1.5090
$N = 480$	M*(0): 8.0866			f(0): 0.0942				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	12.6044	3.6059	2.6283	0.0036	0.0332	0.1259	(0.1668)	2.3140
1	13.0515	3.6773	2.8117	0.0036	0.0337	0.1296	(0.1692)	2.3813
2	12.3891	3.7140	2.5581	0.0037	0.0334	0.1272	(0.1676)	2.3372
3	4.9171	2.7340	0.4840	0.0050	0.0269	0.0843	(0.1350)	1.5490
4	-0.1206	3.4026	0.1773	0.0051	0.0230	0.0624	(0.1155)	1.1461
5	-0.3136	3.6181	0.2017	0.0056	0.0240	0.0688	(0.1209)	1.2635
6	-0.3227	3.6333	0.2035	0.0057	0.0239	0.0684	(0.1204)	1.2564
7	-0.3140	3.5994	0.1996	0.0056	0.0238	0.0673	(0.1195)	1.2368
8	0.0550	3.3089	0.1675	0.0051	0.0228	0.0613	(0.1144)	1.1263
9	2.5597	3.1685	0.2537	0.0066	0.0260	0.0808	(0.1304)	1.4847
ITER.1	1.3475	3.6415	0.2306	0.0068	0.0271	0.0882	(0.0973)	1.6580
ITER.2	1.3336	3.7227	0.2391	0.0067	0.0270	0.0871	(0.0952)	1.6375
ITER.3	2.3780	3.3482	0.2579	0.006	0.0273	0.0888	(0.0943)	1.6678
LOCAL	-0.9520	3.6426	0.2168	0.0142	0.0295	0.1206	(0.1142)	2.2668

Table V $\widehat{M}$, $\hat{f}_{\widehat{M}}(0)$ MODEL 3. AR(1) $\alpha = [0.8]$

$N = 120$	M*(0): 17.4803			f(0): 3.9789				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	-10.6166	2.0697	0.3829	-1.4331	1.4412	0.2609	(0.0121)	0.7774
1	-10.1789	2.1473	0.3542	-1.3518	1.4883	0.2553	(0.0127)	0.7607
2	-10.0005	2.1871	0.3430	-1.3178	1.5076	0.2533	(0.0127)	0.7545
3	-9.5028	2.2363	0.3119	-1.2429	1.5800	0.2553	(0.0141)	0.7605
4	-8.8514	2.2607	0.2731	-1.1201	1.6132	0.2436	(0.0147)	0.7258
5	-9.0624	2.3161	0.2863	-1.1429	1.6180	0.2479	(0.0146)	0.7385
6	-9.0649	2.3162	0.2865	-1.1366	1.6227	0.2479	(0.0148)	0.7386
7	-9.0582	2.3015	0.2859	-1.1340	1.6249	0.2480	(0.0148)	0.7389
8	-9.0917	2.3603	0.2887	-1.1326	1.6280	0.2484	(0.0148)	0.7401
9	-8.9180	2.2433	0.2767	-1.1117	1.6222	0.2443	(0.0147)	0.7278
ITER.1	-9.0712	2.3821	0.2879	-1.1887	1.5021	0.2318	(0.0093)	0.7759
ITER.2	-9.0783	2.3805	0.2883	-1.1870	1.5020	0.2315	(0.0093)	0.7750
ITER.3	-8.9235	2.3304	0.2784	-1.1897	1.5045	0.2324	(0.0094)	0.7779
LOCAL	-9.5462	3.2002	0.3318	-1.2504	1.6060	0.2617	(0.0106)	0.8760
$N = 256$	M*(0): 20.3404			f(0): 3.9789				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	-11.8184	2.3282	0.3507	-1.1649	1.0979	0.1619	(0.0049)	0.9156
1	-11.3429	2.5014	0.3261	-1.1021	1.1276	0.1570	(0.0053)	0.8883
2	-10.9801	2.6084	0.3078	-1.0553	1.1523	0.1542	(0.0054)	0.8724
3	-10.1367	2.8909	0.2686	-0.9537	1.2287	0.1528	(0.0066)	0.8645
4	-8.9645	3.1418	0.2181	-0.7780	1.3120	0.1470	(0.0071)	0.8313
5	-9.2336	3.1412	0.2299	-0.8224	1.3033	0.1500	(0.0072)	0.8486
6	-9.2166	3.1803	0.2298	-0.8212	1.3075	0.1506	(0.0073)	0.8518
7	-9.1907	3.2153	0.2292	-0.8163	1.3121	0.1508	(0.0073)	0.8532
8	-9.1855	3.2708	0.2298	-0.8048	1.3216	0.1512	(0.0073)	0.8555
9	-8.9485	3.0693	0.2163	-0.7688	1.3221	0.1477	(0.0074)	0.8358
ITER.1	-9.1750	3.2453	0.2289	-0.8151	1.3555	0.1580	(0.0087)	0.8343
ITER.2	-9.1983	3.2183	0.2295	-0.8139	1.3549	0.1578	(0.0086)	0.8330
ITER.3	-9.0201	3.2417	0.2221	-0.8117	1.3581	0.1581	(0.0086)	0.8347
LOCAL	-10.4480	4.8921	0.3217	-1.0130	1.5554	0.2176	(0.0105)	1.1489
$N = 480$	M*(0): 23.0653			f(0): 3.9789				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	-13.1144	2.5589	0.3356	-0.9397	0.8959	0.1065	(0.0036)	0.8692
1	-12.5516	2.7019	0.3098	-0.8849	0.9133	0.1021	(0.0038)	0.8338
2	-12.0365	2.9481	0.2887	-0.8150	0.9778	0.1024	(0.0047)	0.8355
3	-11.1065	3.3051	0.2524	-0.7279	1.0481	0.1029	(0.0062)	0.8396
4	-9.2345	3.9710	0.1899	-0.5044	1.2252	0.1109	(0.0114)	0.9052
5	-9.8104	3.8532	0.2088	-0.5864	1.1678	0.1079	(0.0093)	0.8804
6	-9.8532	3.7994	0.2096	-0.5827	1.1720	0.1082	(0.0094)	0.8833
7	-9.8316	3.7731	0.2084	-0.5808	1.1713	0.1080	(0.0094)	0.8813
8	-9.7846	3.7902	0.2070	-0.5720	1.1654	0.1065	(0.0089)	0.8690
9	-9.1792	3.8624	0.1864	-0.5107	1.2173	0.1101	(0.0111)	0.8985
ITER.1	-9.8240	3.7510	0.2079	-0.6154	1.1623	0.1093	(0.0098)	0.8729
ITER.2	-9.8211	3.7482	0.2077	-0.6128	1.1626	0.1091	(0.0098)	0.8716
ITER.3	-9.6628	3.7257	0.2016	-0.6111	1.1682	0.1098	(0.0097)	0.8771
LOCAL	-11.5117	6.4805	0.3280	-0.9551	1.4600	0.1923	(0.0149)	1.5361

Table VI $\widehat{M}$, $\hat{f}_{\widehat{M}}(0)$ MODEL 4. AR(2) $\alpha = [\,1.37, -0.68]$

$N = 120$	M*(0): 9.5192			f(0): 1.7109				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	-1.5246	2.0629	0.0726	0.3645	0.8710	0.3046	(0.0133)	1.0897
1	-1.1855	2.1272	0.0654	0.3387	0.8846	0.3065	(0.0141)	1.0966
2	-1.0976	2.1279	0.0633	0.3331	0.8842	0.3050	(0.0137)	1.0912
3	-1.0675	2.4082	0.0766	0.3115	0.8654	0.2890	(0.0113)	1.0340
4	-6.3775	1.6813	0.4800	0.8904	0.7394	0.4576	(0.0094)	1.6372
5	-5.8612	2.3158	0.4383	0.8097	0.7409	0.4115	(0.0093)	1.4722
6	-6.0061	2.2031	0.4517	0.8097	0.7378	0.4099	(0.0093)	1.4666
7	-6.0687	2.1439	0.4572	0.8149	0.7373	0.4126	(0.0092)	1.4760
8	-6.2126	1.8765	0.4648	0.8163	0.7271	0.4083	(0.0088)	1.4606
9	-1.2837	2.3808	0.0807	0.3197	0.8686	0.2927	(0.0130)	1.0471
ITER.1	-4.0982	3.1207	0.2928	0.5708	0.8716	0.3708	(0.0125)	1.2816
ITER.2	-3.7176	3.2021	0.2657	0.5359	0.8942	0.3713	(0.0131)	1.2831
ITER.3	-1.3912	2.3026	0.0799	0.3027	0.8659	0.2875	(0.0112)	0.9935
LOCAL	-6.8745	1.3891	0.5428	0.6488	0.7208	0.3213	(0.0093)	1.1105
$N = 256$	M*(0): 11.0767			f(0): 1.7109				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	-0.7144	2.4977	0.0550	0.1876	0.6060	0.1375	(0.0031)	1.0817
1	-0.2252	2.6355	0.0570	0.1758	0.6132	0.1390	(0.0029)	1.0939
2	0.0038	2.7185	0.0602	0.1653	0.6149	0.1385	(0.0027)	1.0897
3	0.0732	3.0833	0.0775	0.1629	0.6193	0.1401	(0.0025)	1.1023
4	-8.5512	1.0325	0.6047	0.7154	0.4864	0.2557	(0.0029)	2.0115
5	-8.1785	1.8234	0.5723	0.7207	0.4880	0.2588	(0.0028)	2.0362
6	-8.2247	1.7684	0.5768	0.7123	0.4908	0.2556	(0.0026)	2.0111
7	-8.2768	1.6320	0.5800	0.7096	0.4874	0.2531	(0.0027)	1.9918
8	-8.2397	1.6825	0.5764	0.7049	0.4914	0.2522	(0.0027)	1.9846
9	-0.2085	2.9149	0.0696	0.1636	0.6025	0.1332	(0.0015)	1.0477
ITER.1	-2.2004	4.3069	0.1906	0.2736	0.6748	0.1812	(0.0034)	1.3529
ITER.2	-1.8605	4.1901	0.1713	0.2501	0.6725	0.1759	(0.0034)	1.3133
ITER.3	-0.3725	2.9226	0.0707	0.1635	0.6198	0.1404	(0.0021)	1.0484
LOCAL	-8.5494	1.3252	0.6100	0.5649	0.4803	0.1878	(0.0041)	1.4027
$N = 480$	M*(0): 12.5607			f(0): 1.7109				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	-0.1311	2.6874	0.0459	0.1391	0.4795	0.0852	(0.0034)	1.1124
1	0.3434	2.8332	0.0516	0.1358	0.4871	0.0874	(0.0033)	1.1412
2	0.6247	2.9002	0.0558	0.1269	0.4887	0.0871	(0.0030)	1.1379
3	0.5923	3.2458	0.0690	0.1324	0.4918	0.0886	(0.0029)	1.1574
4	-10.2465	0.7777	0.6693	0.6485	0.3618	0.1884	(0.0043)	2.4608
5	-10.0893	1.1308	0.6533	0.6884	0.3637	0.2071	(0.0043)	2.7052
6	-10.1220	0.9528	0.6551	0.6825	0.3633	0.2042	(0.0044)	2.6678
7	-10.1524	0.8978	0.6584	0.6781	0.3624	0.2020	(0.0044)	2.6380
8	-10.0866	1.0434	0.6518	0.6816	0.3671	0.2048	(0.0044)	2.6747
9	0.4940	2.9411	0.0564	0.1251	0.4836	0.0852	(0.0033)	1.1134
ITER.1	-0.5072	4.3563	0.1219	0.1627	0.5008	0.0947	(0.0056)	1.3540
ITER.2	-0.4021	4.2216	0.1140	0.1464	0.4918	0.0899	(0.0051)	1.2853
ITER.3	0.4426	3.1473	0.0640	0.1255	0.4711	0.081	(0.0048)	1.1605
LOCAL	-9.9984	1.268	0.6438	0.5786	0.4020	0.1696	(0.0045)	2.4236

Table VII $\widehat{M}$, $\widehat{f}_{\widehat{M}}(0)$ MODEL 5. AR(5) α = [0.90, -0.40, 0.30, -0.50, 0.30]

$N = 120$	M*(0): 10.2862			f(0): 0.9947				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	-3.1082	2.4496	0.1480	-0.1328	0.3708	0.1568	(0.0059)	0.7418
1	-3.3923	2.4079	0.1636	-0.1364	0.3588	0.1489	(0.0037)	0.7045
2	-3.5374	2.3912	0.1723	-0.1463	0.3372	0.1366	(0.0018)	0.6462
3	-4.3022	2.2374	0.2222	-0.1631	0.3219	0.1316	(0.0017)	0.6228
4	-6.5136	1.9262	0.4361	-0.0979	0.3082	0.1057	(0.0017)	0.5001
5	-6.3467	2.0432	0.4202	-0.1044	0.3106	0.1085	(0.0059)	0.5135
6	-6.4044	1.9119	0.4222	-0.1079	0.3096	0.1086	(0.0061)	0.5141
7	-6.4625	1.8452	0.4269	-0.1108	0.3017	0.1044	(0.0054)	0.4940
8	-6.4481	1.7389	0.4215	-0.1136	0.2956	0.1014	(0.0056)	0.4797
9	-4.2181	2.2987	0.2181	-0.1474	0.3456	0.1427	(0.0049)	0.6751
ITER.1	-5.6377	2.3226	0.3514	-0.1243	0.3065	0.1106	(0.0028)	0.5476
ITER.2	-5.6777	2.2877	0.3541	-0.1262	0.3062	0.1108	(0.0029)	0.5490
ITER.3	-4.1307	2.3309	0.2126	-0.1602	0.3331	0.1380	(0.0019)	0.6838
LOCAL	-6.7268	2.1374	0.4709	-0.1482	0.3112	0.1200	(0.0053)	0.5947
$N = 256$	M*(0): 11.9693			f(0): 0.9947				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	-1.8053	2.6174	0.0706	-0.0874	0.2974	0.0971	(0.0073)	0.9079
1	-2.3472	2.8392	0.0947	-0.0953	0.2871	0.0925	(0.0071)	0.8646
2	-2.8374	3.0196	0.1198	-0.1056	0.2714	0.0857	(0.0072)	0.8015
3	-4.9642	2.8509	0.2287	-0.1437	0.2331	0.0758	(0.0055)	0.7087
4	-7.9976	2.0997	0.4772	-0.1103	0.2311	0.0663	(0.0042)	0.6196
5	-7.8982	2.0823	0.4657	-0.1119	0.2269	0.0647	(0.0047)	0.6050
6	-7.9106	2.0095	0.4650	-0.1151	0.2188	0.0618	(0.0046)	0.5777
7	-7.9219	1.9699	0.4651	-0.1161	0.2187	0.0619	(0.0046)	0.5793
8	-7.9024	1.8739	0.4604	-0.1193	0.2098	0.0589	(0.0045)	0.5505
9	-4.8077	3.1417	0.2302	-0.1319	0.2579	0.0848	(0.0053)	0.7929
ITER.1	-6.2025	3.2035	0.3402	-0.1221	0.2450	0.0757	(0.0046)	0.7080
ITER.2	-6.1708	3.2312	0.3387	-0.1269	0.2408	0.0749	(0.0046)	0.7001
ITER.3	-4.5963	3.1678	0.2175	-0.1328	0.2584	0.0853	(0.0048)	0.7979
LOCAL	-8.3630	2.4637	0.5306	-0.1619	0.2388	0.0841	(0.0057)	0.7865
$N = 480$	M*(0): 13.5727			f(0): 0.9947				
Method	bias	sd	Rmse	bias	sd	Rmse	(sd)	MSE ratio
GLOBAL	-1.3650	2.6292	0.0476	-0.0555	0.2484	0.0655	(0.0098)	0.9414
1	-1.6950	3.0532	0.0662	-0.0563	0.2460	0.0644	(0.0098)	0.9260
2	-2.2844	3.3601	0.0896	-0.0651	0.2389	0.0620	(0.0094)	0.8917
3	-6.2928	3.2159	0.2711	-0.1293	0.1736	0.0474	(0.0070)	0.6815
4	-9.4970	2.1434	0.5145	-0.1158	0.1638	0.0407	(0.0064)	0.5849
5	-9.4377	2.1885	0.5095	-0.1131	0.1649	0.0404	(0.0064)	0.5811
6	-9.3505	2.1854	0.5005	-0.1131	0.1619	0.0394	(0.0065)	0.5670
7	-9.3790	2.0884	0.5012	-0.1148	0.1570	0.0382	(0.0064)	0.5499
8	-8.8408	2.6853	0.4634	-0.1119	0.1502	0.0354	(0.0061)	0.5099
9	-4.7637	3.9881	0.2095	-0.1081	0.2097	0.0563	(0.0078)	0.8094
ITER.1	-6.0425	4.3909	0.3029	-0.1176	0.2023	0.0553	(0.0043)	0.7600
ITER.2	-6.0598	4.2327	0.2966	-0.1170	0.2027	0.0554	(0.0084)	0.7606
ITER.3	-4.4998	3.7678	0.1870	-0.1189	0.2087	0.0583	(0.0085)	0.8010
LOCAL	-10.0589	2.4984	0.5831	-0.1767	0.1718	0.0614	(0.0071)	0.8430

Table VIII $\widehat{f}(\lambda), \quad \lambda \in [-\pi, \pi]$

Sample size:		$N = 120$		$N = 256$		$N = 480$	
Model	m	IMSE	sd	IMSE	sd	IMSE	sd
2	1	0.13646	(0.09833)	0.06545	(0.03223)	0.03805	(0.01531)
	2	0.15758	(0.11871)	0.07140	(0.04125)	0.03659	(0.01570)
	3	0.13433	(0.09717)	0.06390	(0.03241)	0.03528	(0.01478)
	4	0.13528	(0.09387)	0.06436	(0.03126)	0.03503	(0.01540)
3	1	0.74030	(0.73461)	0.25241	(0.17345)	0.11225	(0.04458)
	2	0.60013	(0.58166)	0.23447	(0.14721)	0.11188	(0.04153)
	3	0.66356	(0.61461)	0.23208	(0.15650)	0.09898	(0.03938)
	4	0.66240	(0.61012)	0.22940	(0.15911)	0.09612	(0.03824)
5	1	0.20289	(0.16848)	0.09384	(0.08463)	0.04196	(0.02241)
	2	0.14133	(0.09355)	0.05981	(0.03302)	0.03116	(0.01396)
	3	0.12857	(0.08317)	0.05855	(0.03205)	0.03082	(0.01405)
	4	0.13006	(0.08407)	0.06068	(0.03518)	0.03085	(0.01412)
6	1	0.47824	(0.74073)	0.14747	(0.21982)	0.06533	(0.05338)
	2	0.31678	(0.38975)	0.13110	(0.09898)	0.06334	(0.04976)
	3	0.29050	(0.39069)	0.12113	(0.09192)	0.05858	(0.04625)
	4	0.28964	(0.39775)	0.14747	(0.21982)	0.05835	(0.04717)
7	1	0.13666	(0.09507)	0.09814	(0.08571)	0.04842	(0.01873)
	2	0.15576	(0.11148)	0.10457	(0.05167)	0.05267	(0.02194)
	3	0.13401	(0.09333)	0.09938	(0.04939)	0.05180	(0.01943)
	4	0.13444	(0.09171)	0.09809	(0.04800)	0.04991	(0.01980)

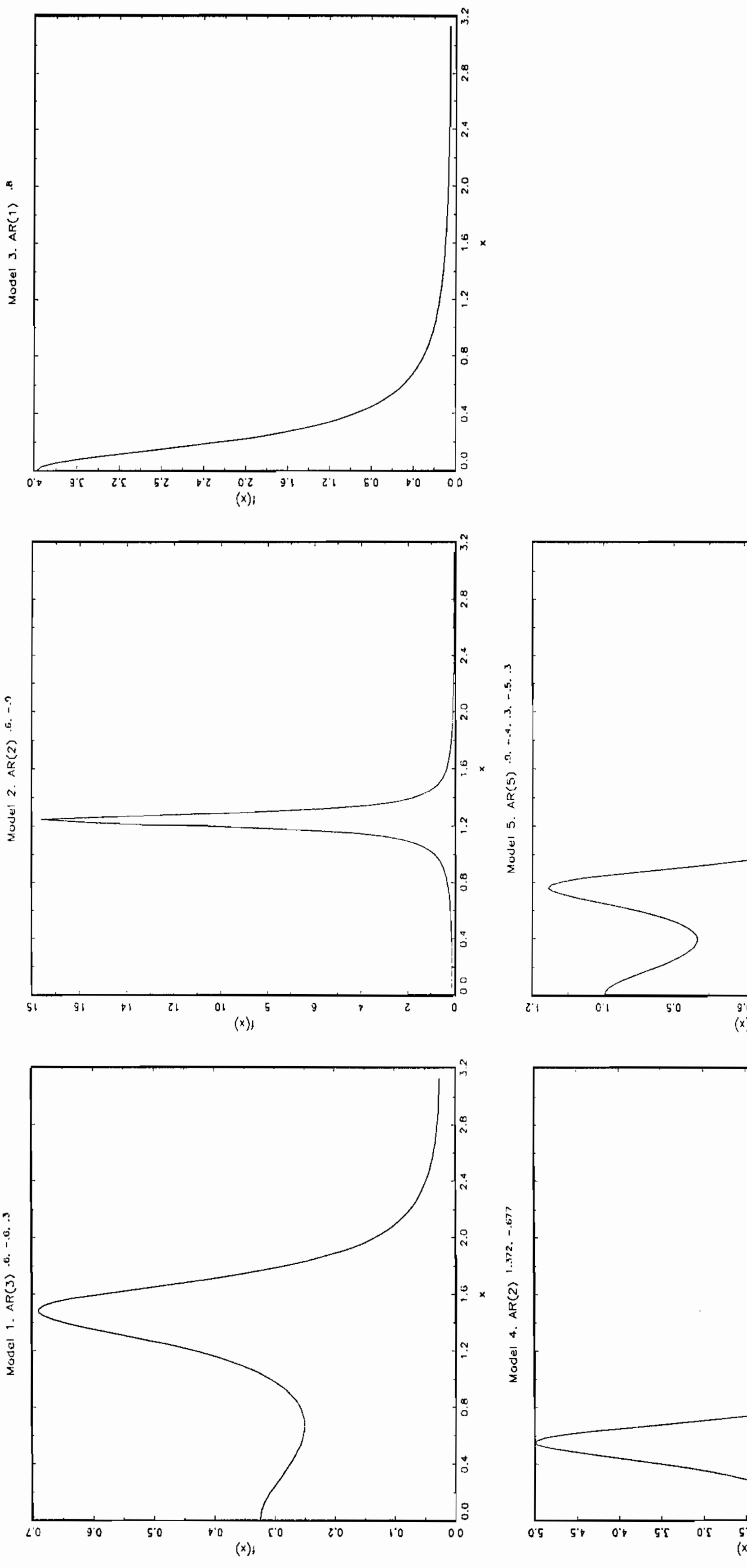


Figure 1: Spectral densities for Models 1 to 5.

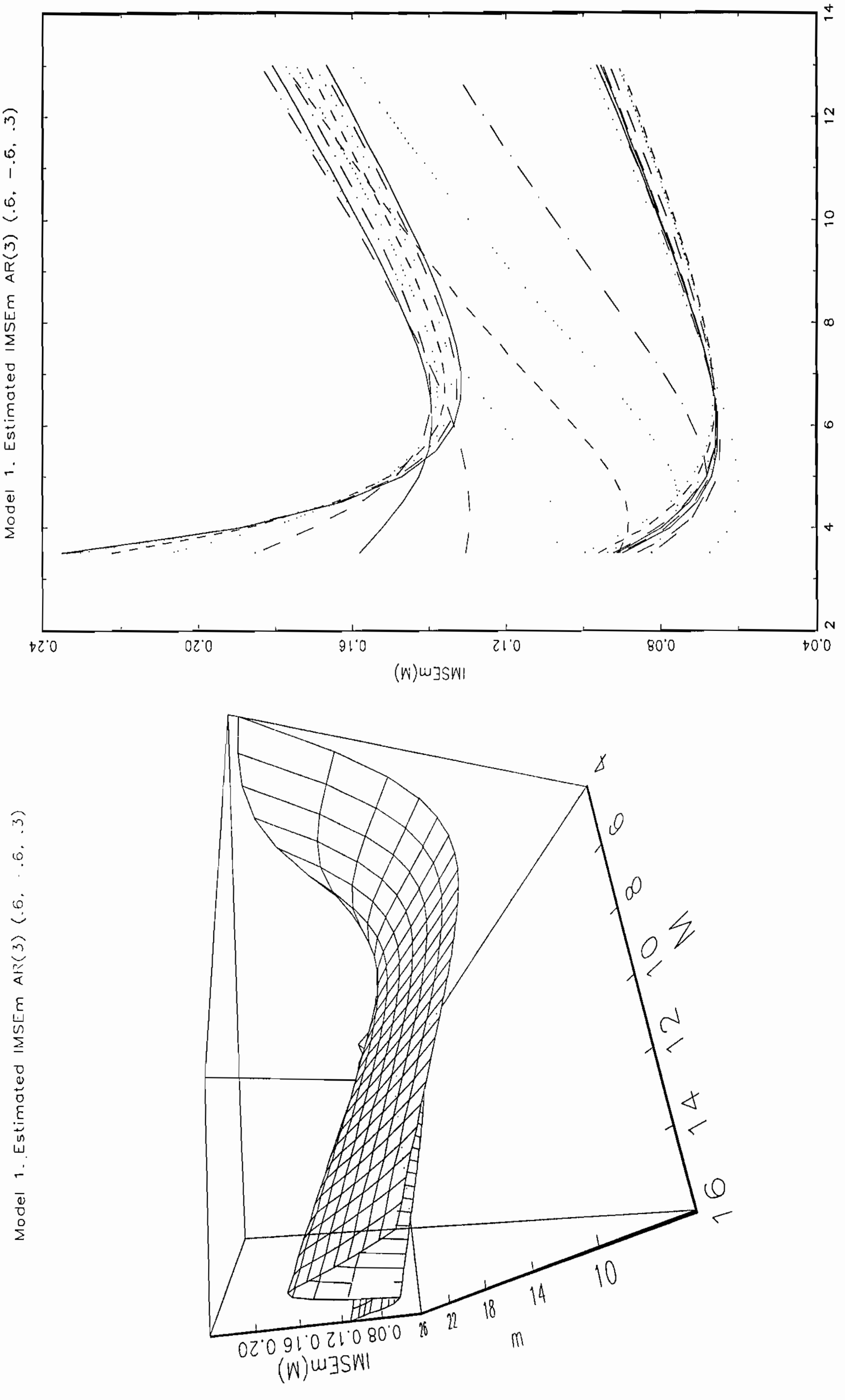
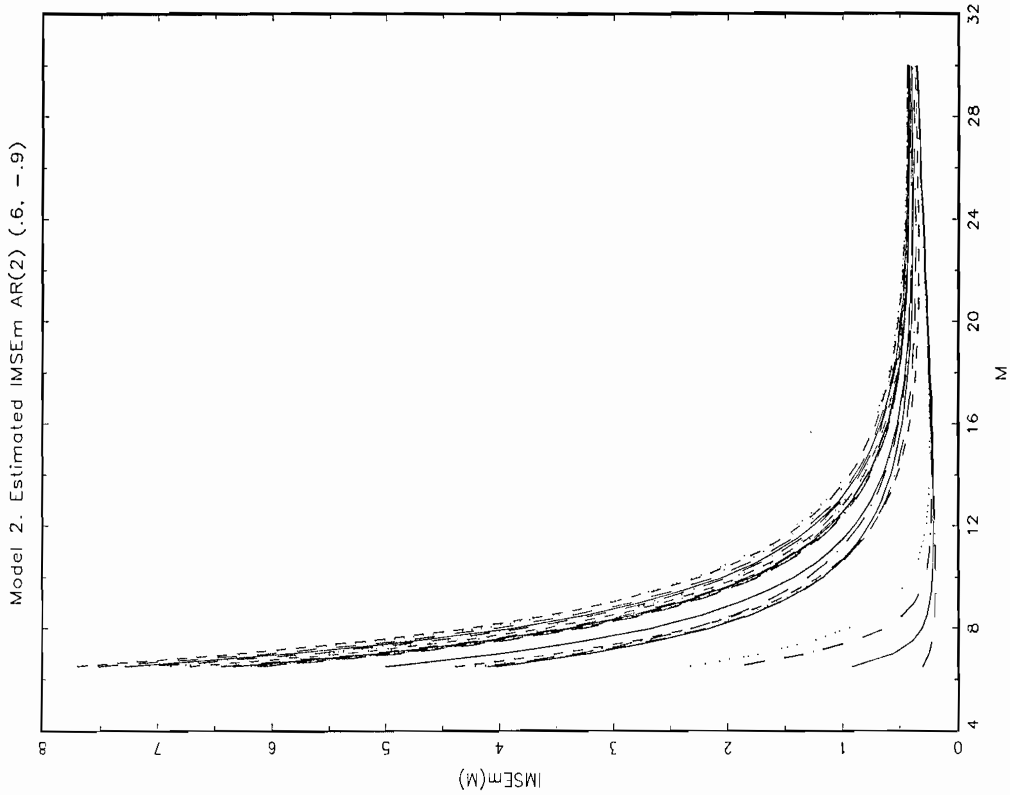


Figure 2: IMSE_m(M,0) for Model 1.



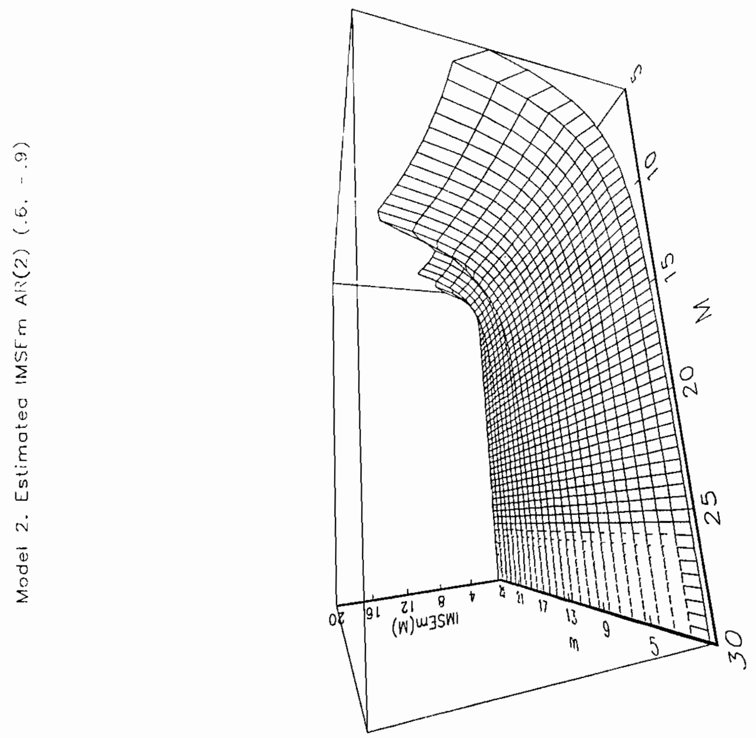


Figure 3: $\text{IMSE}_m(\text{M},0)$ for Model 2.

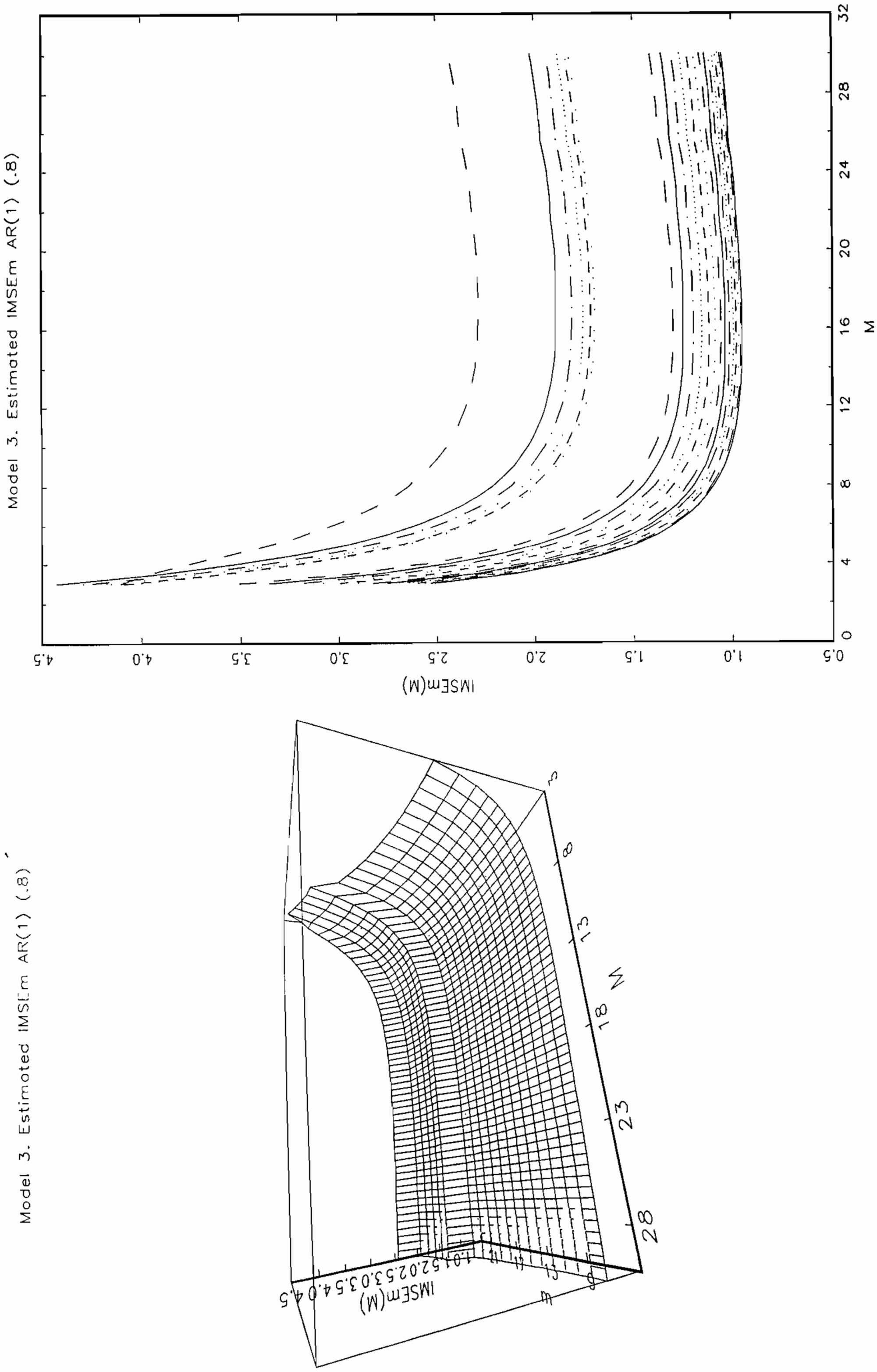
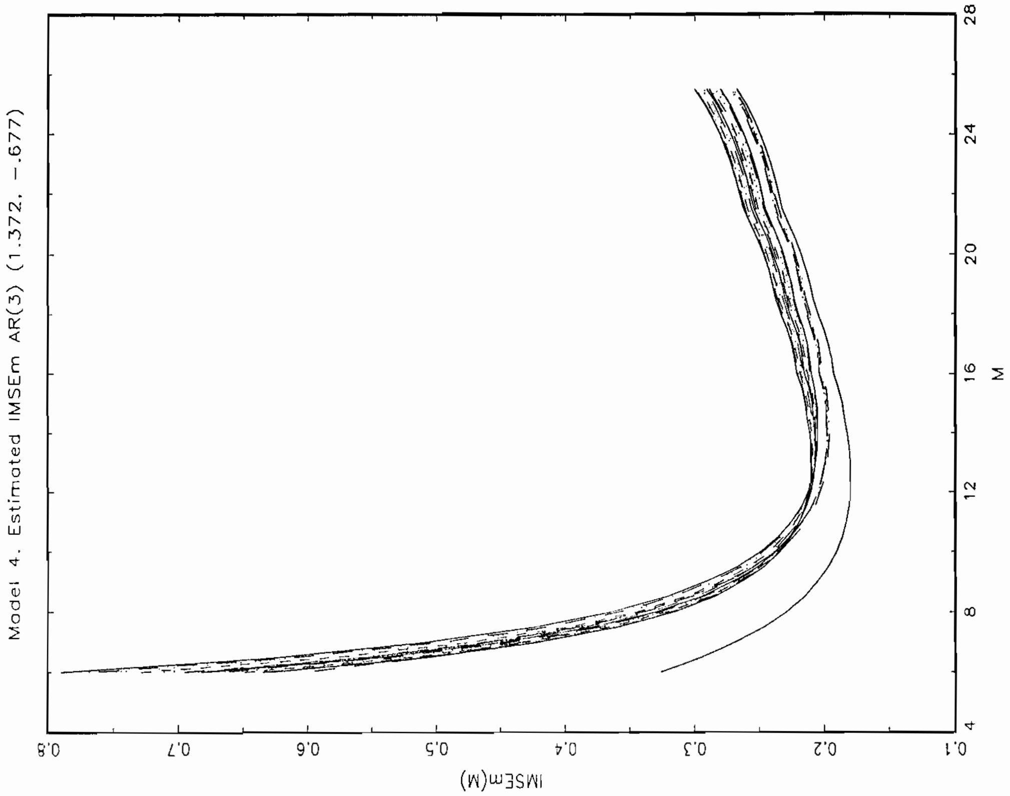


Figure 4: IMSE$_m$(M,0) for Model 3.



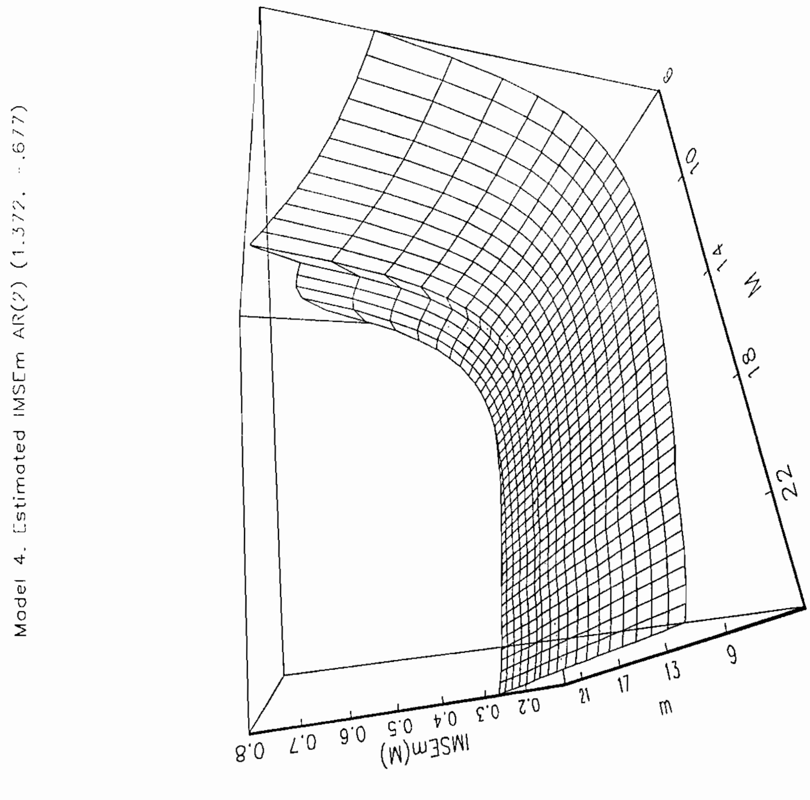
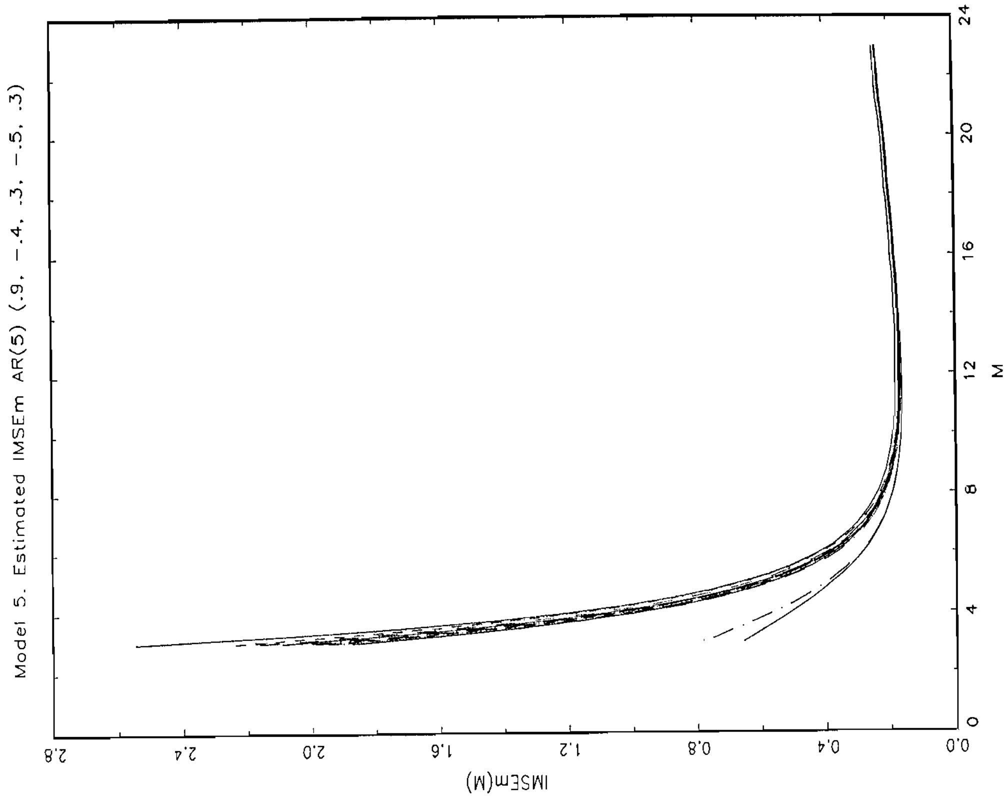


Figure 5: IMSE_m(M,0) for Model 4.



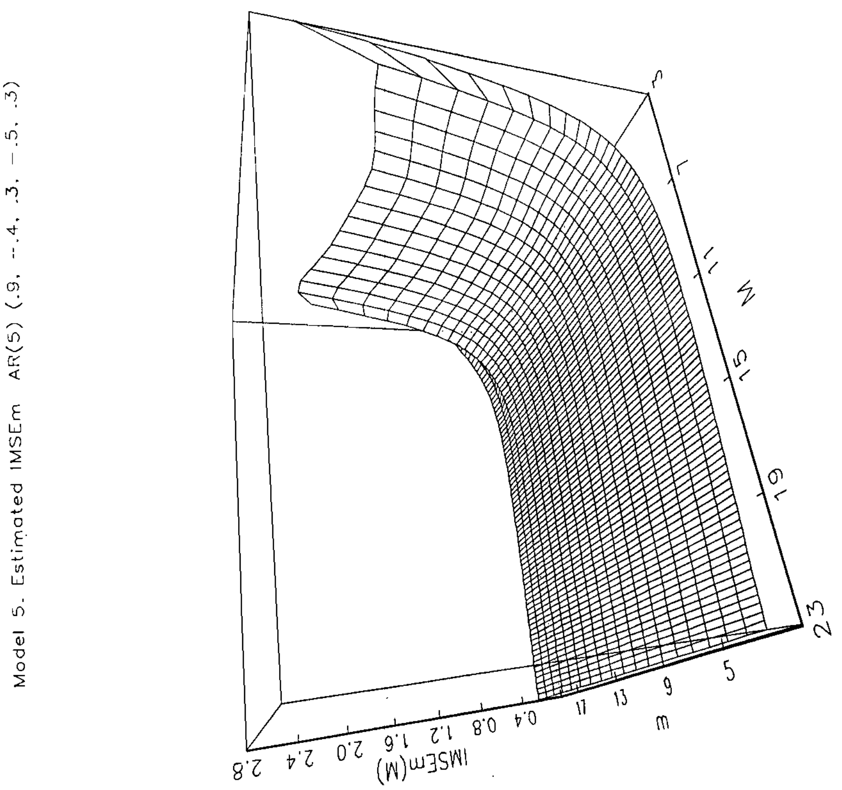


Figure 6: IMSE$_m$(M,0) for Model 5.

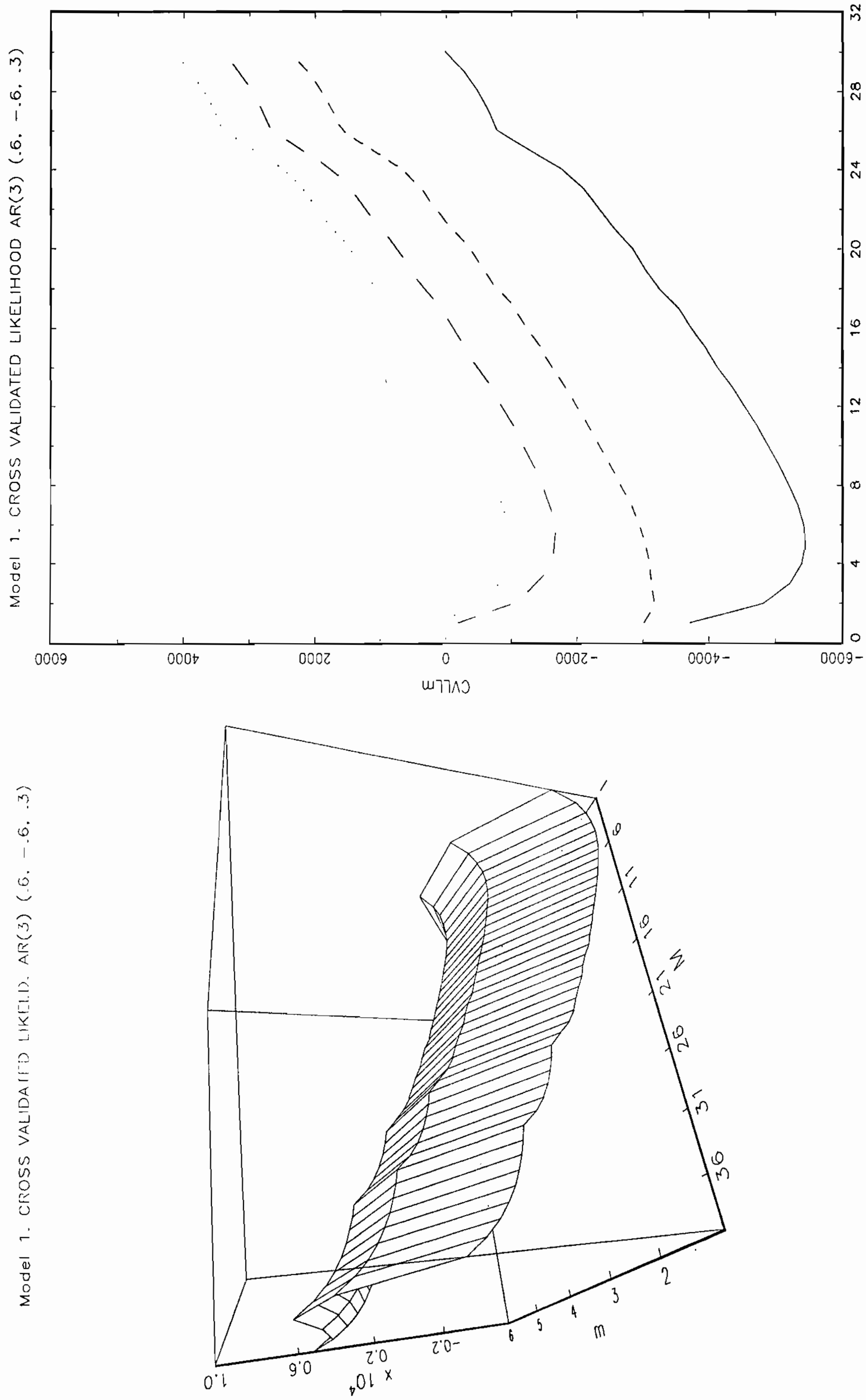


Figure 7: CVLL_m(M,0) for Model 1.

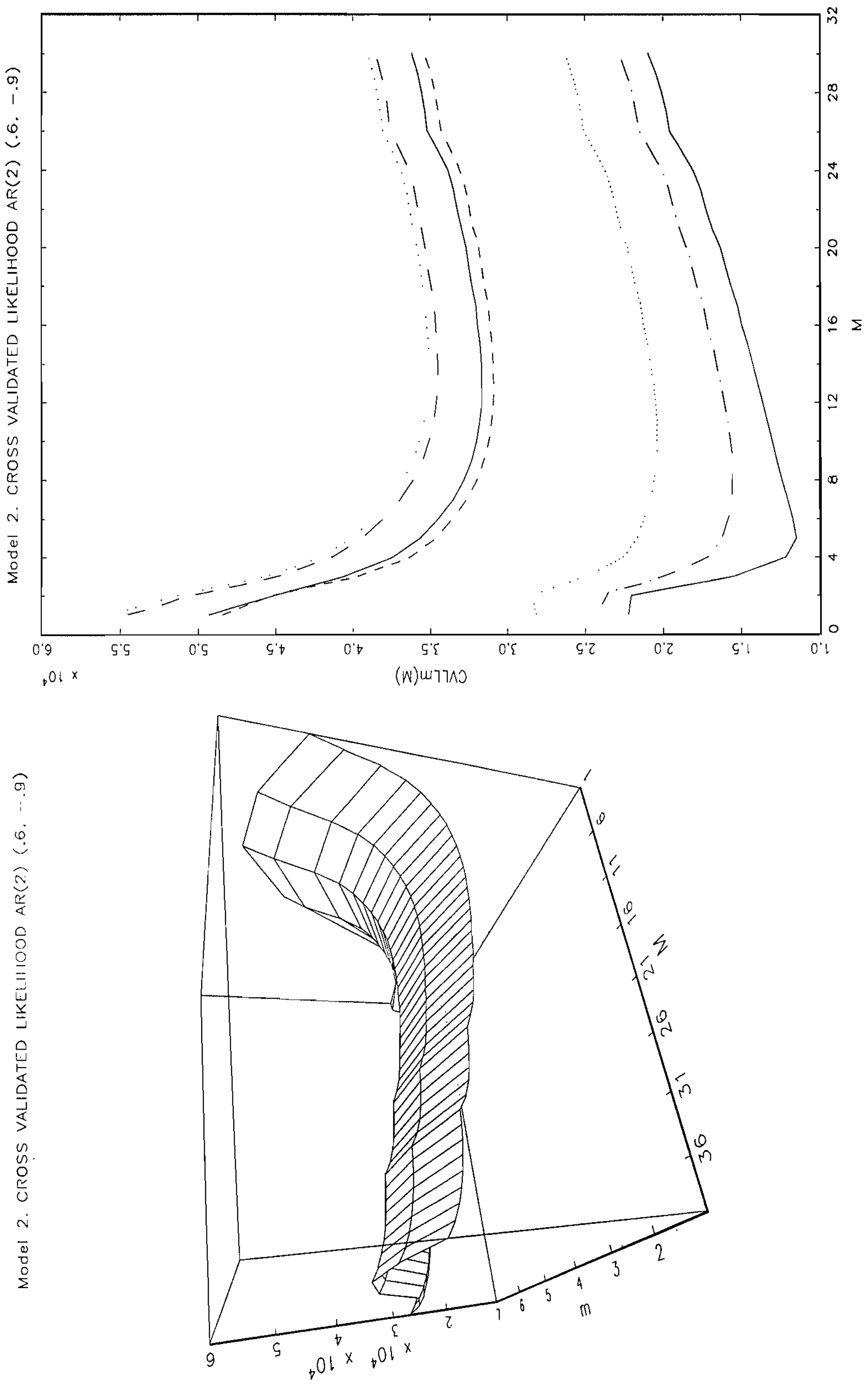


Figure 8: CVLL$_m$(M,0) for Model 2.

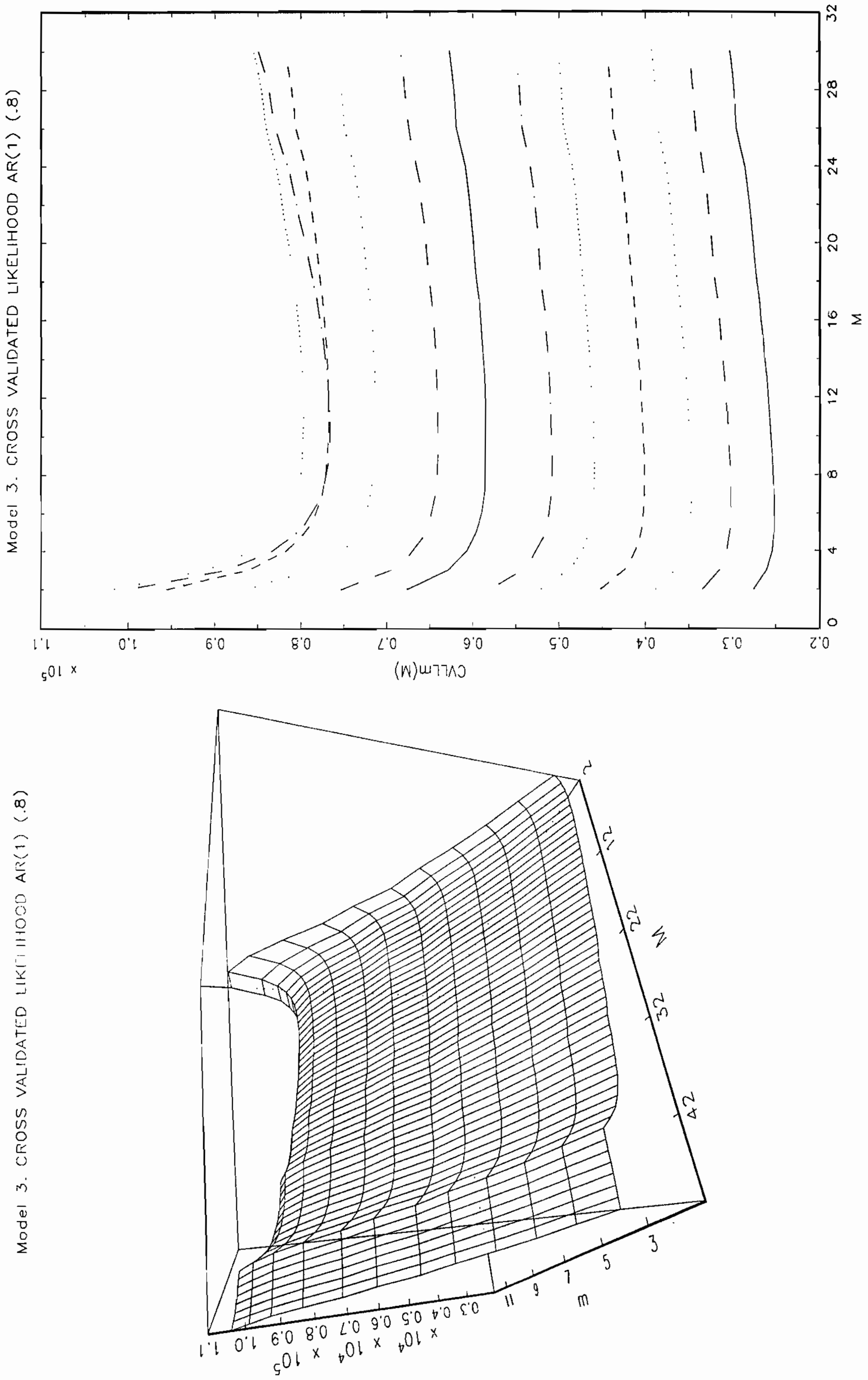
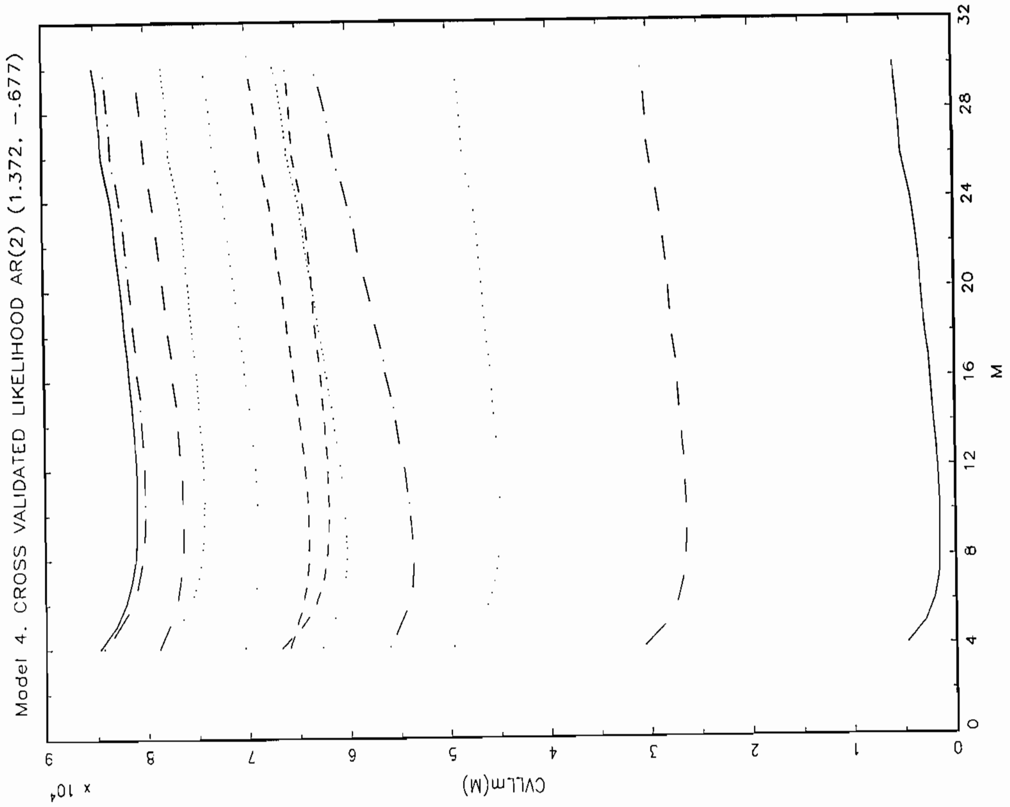


Figure 9: CVLL_m(M,0) for Model 3.



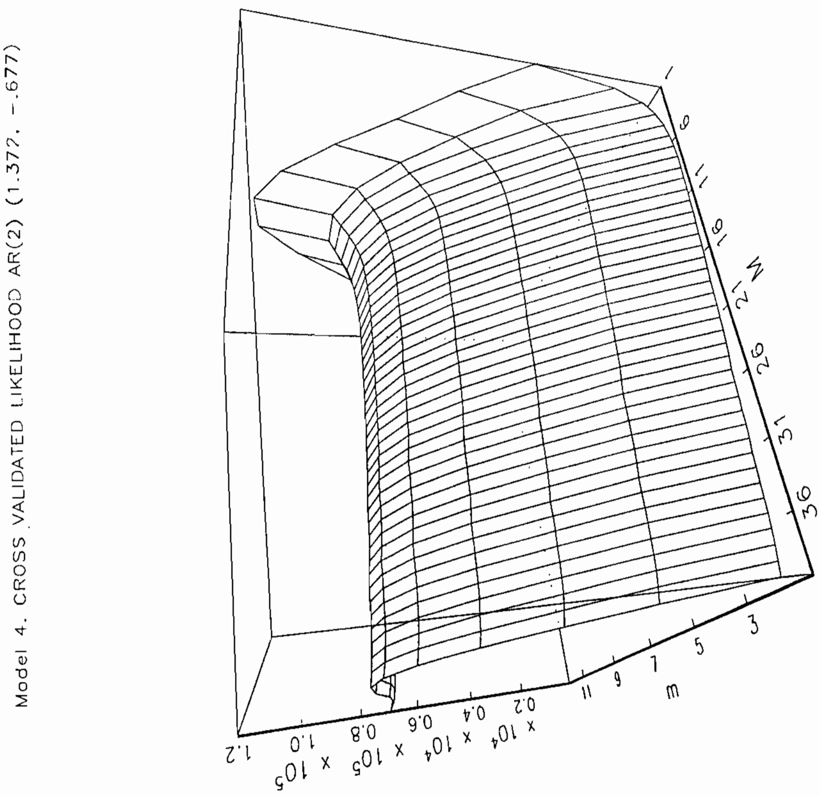
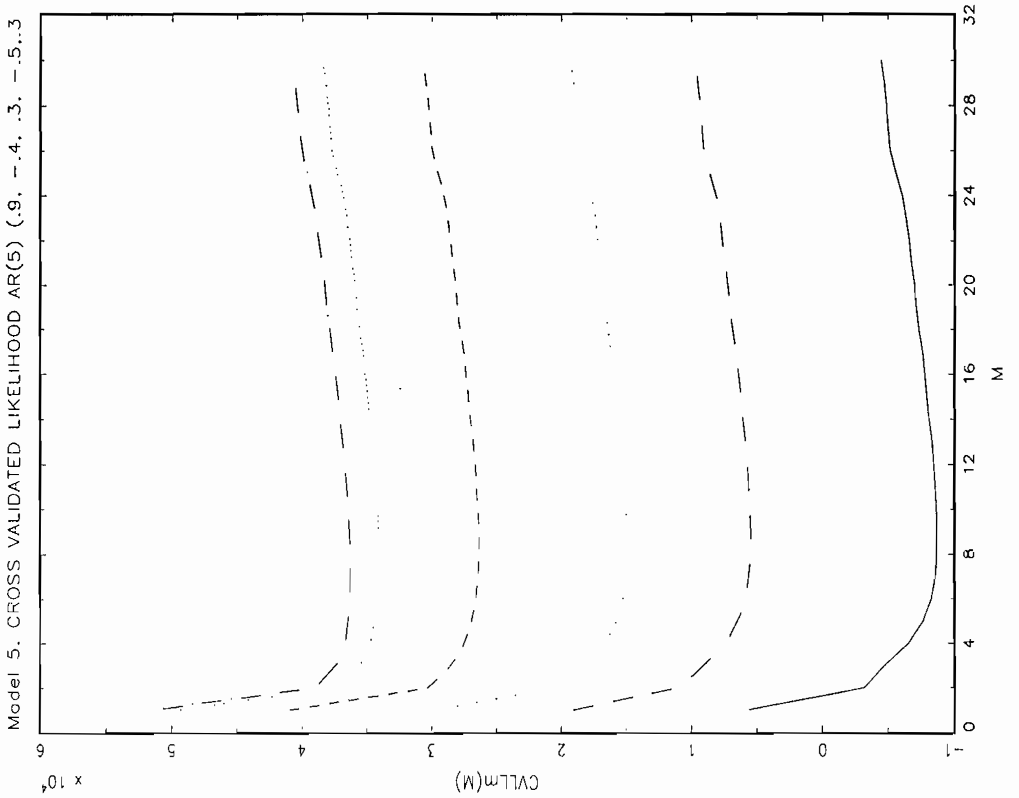


Figure 10: CVLL$_m$(M,0) for Model 4.



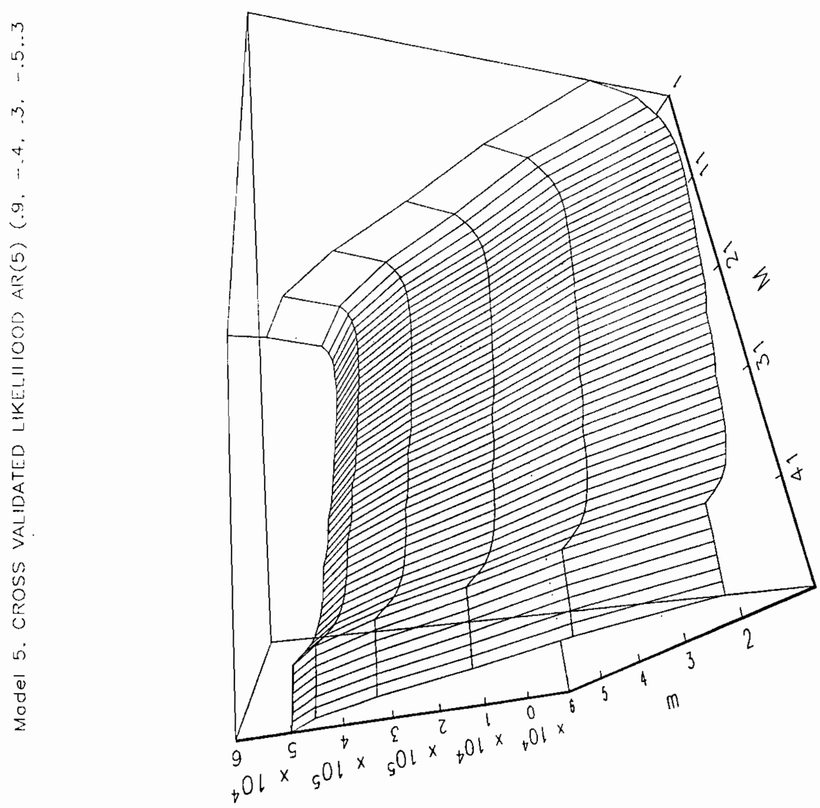


Figure 11: $\mathrm{CVLL}_m(\mathrm{M},0)$ for Model 5.